**University of Alberta**

# Characterization of the BaeSR regulon in *Escherichia coli*

by

Christopher Wilson Oates

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

in

Microbiology and Biotechnology

Department of Biological Sciences

Edmonton, Alberta

Fall 2006

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*
*ISBN: 978-0-494-22336-9*
*Our file Notre référence*
*ISBN: 978-0-494-22336-9*

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

**Abstract**

In *Escherichia coli* the BaeSR pathway responds to chemicals which damage the envelope. The regulon of the BaeSR two-component system is not well characterized, but is known to contain multidrug efflux pumps, MdtABCD and AcrD, and periplasmic protein Spy. This project was intended to identify additional genes in the BaeSR regulon by means of a genetic screen utilizing the λp*lac*Mu53 phage. We identified three insertions which resulted in *lacZ* fusions that were upregulated by BaeR overexpression. These insertions occured in *mdtA, yggC,* and *yliA*. Two insertions, in *fecD* and near *yjgX*, were found to be downregulated by BaeR overexpression. These *lacZ* fusions were transduced into a background containing *baeS1::*Tn*10cam,* a gain of function mutation which upregulates the Bae pathway, resulting in a similar phenotype to that seen during BaeR overexpression. Except for the *mdtA-lacZ* fusion, we were unable to induce expression using chemical activators of the BaeSR pathway. We also examined the ability of BaeR to bind the promoters of the genes identified, demonstrating binding to the promoter of the operon containing *yggC.* Additional work demonstrated MBP-BaeR binding to the promoters of *ygcL* and *yicO,* genes identified in the literature as putatively BaeR regulated, and identified the sequence recognized by MBP-BaeR in the *acrD* promoter. These results present evidence for the inclusion of *yggC*, *yliA*, *fecD* and *yjgX* in the BaeR regulon.

**Acknowledgements**

The past three years I have spent performing research in the Raivio Lab have been an enjoyable period in my life, and have provided me tremendous opportunities for personal and scholarly growth. I wish to sincerely thank my supervisor Dr. Tracy Ravio for her help and support throughout this time, and for allowing me this opportunity. I also wish to thank all of the members of my committee, Drs. Susan Jensen, Brenda Leskiw, and Doug Storey for their helpful suggestions over the course of this thesis.

I also wish to thank all of the members of the Raivio lab, in no particular order, (or rather a very particular, though neutral, order) Daelynn Buelow, Dawn MacRitchie Rebecca Mulamba, Anna Nevesinjac and Nancy Price, for their helpful contributions and discussions both on and off scientific topics, as well as making the lab a very pleasant place to work. I also wish to extend special thanks to former lab members Jordan Ward and Lindsey Kurach, for their friendship, conversation, and for allowing me to maintain whatever tenuous claim I have to sanity at different parts of this project. I will always hold memories of the infamous safety mishaps (and much exaggerated non-existent scarring), endless discussions about politics, the appropriateness of household pets as food sources, and the (non)existence of a purported childhood television icon, as well as all of late night Deep Dish gorging, outings to rock climb, see movies/games, or go curling etc. as fond reminders of my time here.

Finally I wish to thank my family, my parents, brother and sister, for their love and support during the completion of this thesis, and throughout my whole life.

## Thesis Outline


**1. Introduction** 1

1.1 Two-Component Systems 2

1.1.1 Characteristics of Two-Component Systems 2

1.1.1.1 Characteristics of Histidine Kinase Proteins 6

1.1.1.2 Characteristics of Response Regulator Proteins 10

1.2 Envelope Stress Response 15

1.2.1 Characteristics of the Gram Negative Envelope 15

1.2.2 General Characteristics of the Envelope Stress Response 21

1.2.2.1 The $\sigma^E$ Envelope Stress Response Pathway 21

1.2.2.2 The Cpx Envelope Stress Response Pathway 26

1.3 The BaeSR Two-Component System 31

1.3.1 The initial identification of BaeSR 31

1.3.2 The BaeSR Regulon 34

1.3.2.1 BaeSR is involved in the envelope stress response 34

1.3.2.2 BaeSR is involved in Multidrug Resistance 36

1.3.2.3 Other possible functions of the BaeSR pathway 38

1.3.2.4 Other putative regulon members 39

1.4 Multidrug Resistance 40

1.4.1 General Characteristics of Multidrug Efflux Pumps 40

1.4.1.1 Structure and Function of Multidrug Efflux Pumps 42

1.4.1.2 The Regulation of Multidrug efflux 49

1.5 Proposed Research Project 51

1.5.1 A screen to identify BaeSR regulated genes 51

1.5.2 Concluding Remarks 54

**2. Materials and Methods** **55**

2.1 Bacterial Strains, Phage, Plasmids, and Growth Conditions 55

2.1.1 *Escherichia coli* strains used and created in this study 55

2.1.2 Growth and maintenance of *E.coli* strains 55

2.1.3 Phage used in this study 55

2.1.4 Phage preparation and quantification 60

2.1.4.1 λ Phage plaque assay 60

2.1.4.2 High Titre λ liquid lysates 60

2.1.5 Plasmids and Cloning vectors used and created in this study 61

2.2 Transduction and transformation 63

2.2.1 Preparation of *E.coli* competent cells 63

2.2.1.1 Electroporation competent cells 63

2.2.1.2 Chemically competent cells 63

2.2.2 Transformation of *E. coli* cells 64

2.2.2.1 Electroporation 64

2.2.2.2 Heat Shock 64

2.2.3 P1*vir* transduction 65

2.2.3.1 Preparation of P1*vir* lysates 65

2.2.3.2 P1*vir* transductions 65

2.3 DNA Isolation, Analysis and Purification 66

2.3.1 Isolation of plasmid DNA from *E.coli* 66

2.3.2 Polymerase Chain Reaction 66

2.3.2.1 PCR utilizing random primers 73

2.3.3 Agarose gel electroporesis 73

2.3.4 Purification of DNA from agarose gels 74

2.3.5 DNA sequencing 74

2.3.6 Digestion and cloning of DNA 75

2.4 Genetic Screen for BaeSR regulated genes 75

2.4.1 λplacMu53 infection 75

2.4.2 Isolation and screening of transposition mutants 76

2.4.3 Confirmation of a BaeR inducible phenotype 77

2.4.4 Identification of target genes 77

2.5 MBP-BaeR Protein purification and analysis 78

2.5.1 Preparation of whole cell lysates 78

2.5.2 MBP-BaeR purification using affinity chromatography 79

2.5.3 Bradford assay for protein concentration 80

2.5.4 Protein buffer exchange 80

2.5.5 SDS-Polyacrylimide Gel Electrophoresis for protein analysis 80

2.6 Electrophoretic mobility shift assays 80

2.6.1 Preparation of end-labeled DNA fragments 80

2.6.2 Phosphorylation of MBP-BaeR by small phosphor-donor 81

2.6.3 Electrophoretic mobility shift assay 81

2.6.4 Autoradiography 82

2.7 DNaseI Protection assay 83

2.7.1 Preparation of end-labeled DNA fragments 83

2.7.2 Purification of radio-labled DNA from polyacrylimide gels 83

2.7.3 DNaseI Protection assay 84

2.7.4 Sanger sequencing protocol 85

2.7.5 Polyacrylimide gel electrophoresis for sequencing 85

2.8 Reporter assays 86

2.8.1 β-galactosidase assay 86

**3. Characterization of Novel BaeSR regulated genes 88**

3.1 Genes identified by transposon screen 88

3.2 Induction of operon fusions by the *baeSI*::Tn*10cam* allele 100

3.3 Induction of operon fusions via inducers of the BaeSR pathway 108

**4. Molecular Characterization of the BaeSR regulon 117**

4.1 BaeR binding to promoter regions of putative regulon members 117

4.2 Identification of putative BaeR binding recognition sequences 125

4.3 DNaseI protection assay of the *acrD* promoter 127

**5. Discussion 134**

5.1 Diversity of the BaeSR regulon 134

5.1.1 Novel genes identified in this study 134

5.1.1.1 *yliA* 135

5.1.1.2 *yjgX* 137

5.1.1.3 *yggC* 138

5.1.1.4 *fecD* 140

5.1.2 Other members of the BaeSR regulon 142

5.1.2.1 *yicO, ygcL,* and *ycaC* 142

5.1.2.2 *mdtA, acrD,* and *spy* 143

5.2 Regulation by BaeSR 144

5.2.1 The BaeR consensus sequence 144

5.2.2 Mechanism of BaeR regulation 147

5.2.3 BaeSR activating signals 147

5.2.3.1 *baeS1*::Tn*10cam* 148

5.2.3.2 Chemical activating signals 150

5.3 Future Directions 153

**6. References** **160**

**List of Tables**

**Chapter 2:**


| | | |
|---|---|---|
| **2.1** | *Escherichia coli* strains | 56 |
| **2.2** | Phage Strains | 59 |
| **2.3** | Plasmids and Cloning vectors | 62 |
| **2.4** | Oligonucleotide Primers | 67 |


**Chapter 4:**


| | | |
|---|---|---|
| **4.1** | Putative BaeR recognition sequences | 126 |

## List of Figures

**Chapter 1:**


**1.1** Signal transduction in a model two-component system. 4

**1.2** The Gram-negative envelope. 17

**1.3** The $\sigma^E$ envelope stress response 23

**1.4** The CpxAR envelope stress response. 28

**1.5** The BaeSR envelope stress response. 32

**1.6** The AcrAB multidrug efflux pump. 43


**Chapter 3:**


**3.1** Schematic diagram of λp*lac*Mu53 insertion in the bacterial chromosome. 89

**3.2** λp*lac*Mu53 insertions in *mdtA, yggC* and *yliA* upregulated by BaeR overexpression. 93

**3.3** λp*lac*Mu53 insertions in *fecD* and near *yjgX* downregulated by BaeR overexpression. 95

**3.4** Chromosomal insertion points of the λp*lac*Mu53 mutations. 97

**3.5** λp*lac*Mu53 insertions in *mdtA, yggC and yliA* activated by the *baeS1::*Tn*10cam* allele. 101

**3.6** λp*lac*Mu53 insertions in *fecD* and near *yjgX,* downregulated by the *baeS1::*Tn*10cam* allele. 104

**3.7** β-galactosidase assay showing the affect of the *baeS1::*Tn*10cam* allele upon expression of β–galactosidase by the λp*lac*Mu53 insertion mutations. 106

**3.8** λp*lac*Mu53 insertions in *mdtA, yggC and yliA* in a Δ*baeR* background 110

**3.9** λp*lac*Mu53 insertions in *fecD* and near *yjgX* in a Δ*baeR* background 112

**3.10** Induction of β–galactosidase in strains carrying a λp*lac*Mu53 insertion in *mdtA* by inducers of the BaeSR pathway. 115

**Chapter 4:**

**4.1** Locations chosen for EMSA study 118

**4.2** Electrophoretic mobility shift assays demonstrating BaeR binding to the *acrD, ygcL, yicO,* and *cmtB* promoter regions. 121

**4.3** DNaseI protection assay of the *acrD* promoter coding strand 129

**4.4** Alignment of the BaeR binding sites in the *acrD* and *mdtA* promoters 131

**Chapter 5:**

**5.1** A revised view of the BaeSR pathway 157

**List of Abbreviations**

| | |
|---|---|
| A | Absorbance or adenosine |
| ABC | ATP binding cassette |
| Amp | Ampicillin |
| APS | Ammonium Persulfate |
| Ara | Arabinose |
| ATP | Adenosine-triphosphate |
| bp | Base pairs |
| C | Cytosine |
| Cam | Chloramphenicol |
| C-terminal | Carboxy terminal |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| ECF | Extracytoplasmic factor |
| EDTA | ethylenediaminetetraacetic acid |
| EMSA | Electrophoretic mobility shift assay |
| G | Guanine |
| xG | times Gravity |
| HK | Histidine kinase |
| IPTG | Isopropyl-B-D-galactopyranoside |
| IM | Inner Membrane |
| Kan | Kanamycin |
| kb | Kilobase pair |

| | |
|---|---|
| L | Litre |
| LB | Luria Broth |
| Lux | Luciferase |
| M | Molar |
| MBP | Maltose binding protein |
| MF | Major facilitator |
| MFP | Membrane fusion protein |
| mg | Milligram |
| mL | Millilitre |
| mM | Millimolar |
| N-terminal | Amino-terminal |
| OD600 | Optical Density |
| OM | Outer membrane |
| O/N | Overnight |
| P | Phosphate |
| PBS | Phospho-buffered saline |
| PCR | Polymerase chain reaction |
| pmol | Picomoles |
| PP | Periplasm |
| PTS | Phosphoenol pryuvate: carbohydrate phosphotransferase system |
| RND | Resistance/Nodulation/Cell division |
| rpm | Revolutions per minute |
| RR | Response regulator |

| | |
|---|---|
| RT | Room temperature , real time (-PCR) |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SIC | Single isolated colony |
| Spc | Spectinomycin |
| T | Thymine |
| TAE | Tris-acetate EDTA |
| TBE | Tris-borate EDTA |
| Tris | (hydroxymethyl) aminomethane |
| U | Units |
| V | Volts |
| W | Watts |
| WT | Wild type |
| w/v | weight per volume |
| α | Alpha |
| β | Beta |
| Δ | Deletion |
| σ | Sigma |
| °C | Degrees Celsius |
| µg | microgram |
| µL | microlitre |
| µM | micromolar |

# 1. Introduction

Bacteria live in a dynamic environment in which constant fluctuations in environmental conditions may threaten the survival of the cell. In order to meet these environmental challenges, bacteria must be able to sense changes in external conditions and modify their gene expression in an adaptive response to the stresses they face. One of the major mechanisms utilized by bacteria in order to achieve this is two-component regulatory systems, which consist of two proteins, a histidine hinase, and a cognate response regulator. Together these proteins are able to detect information about extracellular conditions, and transduce this information across the cell envelope in order to affect a cellular response via an alteration of gene expression.

Among the major classes of environmental stresses that are encountered by bacterial cells is envelope stress. As the outermost component of the cell, the envelope is the site of contact with hazardous environmental conditions, and as such must be maintained if the cell is to survive. The envelope stress response in *Escherichia coli* is known to be controlled by three overlapping systems, the alternate sigma factor $\sigma^E$, the two-component system CpxAR and the recently identified BaeSR two-component system. Together these three systems control the expression of a variety of proteins involved in the folding, stability and degradation of envelope proteins.

Another stress presented to bacteria both in the environment and in a clinical setting is the presence of antibiotics and other toxic compounds. Antibiotics are substances which are able to interfere with the normal functioning of the cell, resulting in either death or the prevention of cellular division. Resistance to a wide variety of

antibiotics is conferred by the expression of membrane associated proteins known as multidrug efflux pumps. Multidrug efflux pumps allow for the export of these compounds across both membranes of the Gram-negative envelope and back into the external environment. Despite this ability, both the control, true physiological function and native substrates of these systems are not clear, although a large body of knowledge is available on the subject.

The BaeSR two-component system has been less studied, and as a result of this, an understanding of its role in the physiology of *E. coli* is only beginning to emerge. It is believed to constitute a new envelope stress response system, and is known to control the expression of two multidrug exporters.

This thesis presents a transpositional mutagenesis screen to further characterize the BaeSR regulon, and subsequent molecular genetic investigations of the genes identified using electrophoretic mobility shift assays, and β-galactosidase assays in a variety of genetic backgrounds.

## 1.1 Two-Component Systems

### 1.1.1 Characteristics of Two-Component Systems.

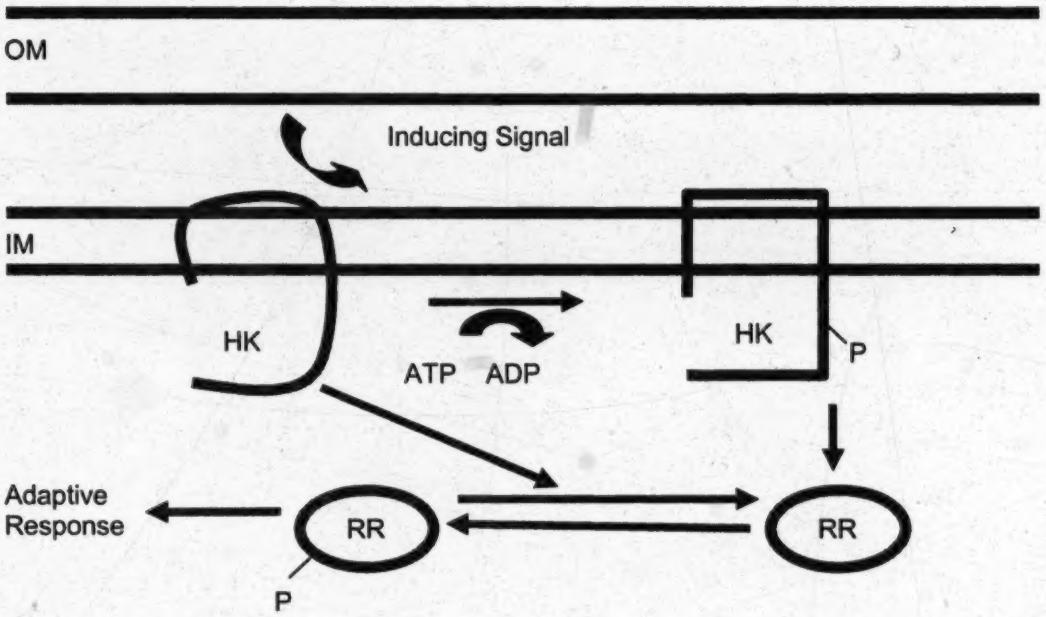
In order to adapt to potentially threatening environmental conditions, bacteria must first be able to sense them, and then must alter their gene expression in a way that responds specifically to the stress they are facing. In many bacteria, this is done by two-component regulatory systems. The archetypical two-component system consists of a sensor histidine kinase protein and a response regulator protein (Parkinson 1993; Hoch

and Silhavy 1995; Appleby *et al.* 1996; Raivio and Silhavy 2001a; West and Stock 2001). These protein pairs are able to detect specific environmental signals and alter gene expression accordingly. This transduction of information is achieved through a phospho-relay system, in which autophosphorylation of a conserved histidine residue in the histidine kinase occurs upon detection of the activating signal, and the phosphate moiety is transferred to a conserved aspartate residue in the response regulator (Parkinson 1993; Hoch and Silhavy 1995; Appleby *et al.* 1996; Raivio and Silhavy 2001a; West and Stock 2001) (Figure 1.1). This allows the response regulator to function as a transcriptional regulator, altering the expression of a subset of genes involved in dealing with the environmental stress being encountered, the regulon of the system.

This conservation of the phospho-relay mechanism amongst diverse two-component regulatory systems is reflected in a conservation of the sequences of the genes that encode them. It was this widespread conservation of motifs in both histidine kinases and response regulators which first lead to their identification as a common system for regulating gene expression in response to environmental signals (Nixon *et al.* 1986). It was noted that the NtrB and NtrC proteins of *Bradyrhizobium parasponia*, which are involved in the response to nitrogen limiting conditions, share homology with an extensive range of regulatory proteins in diverse bacterial species. Specifically, carboxyl-terminal (C-terminal) homology was noted between histidine kinases NtrB of *B. parasponia*, EnvZ, CheA and CpxA of *E. coli*, and VirA of *Agrobacterium tumefaciens*. Furthermore, amino-terminal (N-terminal) homology was seen between the response regulators NtrC of *B. parasponia*, OmpR, PhoB, CheY and SfrA of *E. coli*, and VirG of *A. tumefaciens* (Nixon *et al.* 1986). This model for response to environmental signals also

**Figure 1.1. Signal transduction in a model two-component system.**

The presence of an inducing cue activates the HK as an autokinase. This phosphate is then transferred to the RR, activating it as a regulator of transcriptional activity. When inducing cues are absent, the HK no longer functions as a autokinase, and instead acts as a phosphatase, dephosphorylating the RR, and returning the system to its resting state.


OM
Inducing Signal
IM
HK
ATP
ADP
HK
P
Adaptive
Response
RR
RR
P

suggested that differences in the unique sequences between the above proteins allows for them to detect and respond to a wide variety of differing signals. More specifically, variation in the histidine kinase N-terminus would allow detection of diverse activating cues, whereas the diversity of the response regulator C-terminus would allow the activation of distinct subsets of genes. The conservation of the histidine kinase C-terminal and the response regulator N-terminal reflect conservation of the autophosphorylation and phosphate receiver domains respectively. This conservation of functional domains, allows us to suppose an ancestral protein pair which diversified through evolution to control a diverse selection of cellular responses.

#### 1.1.1.1 Characteristics of Histidine Kinase (HK) Proteins.

Histidine kinase proteins serve as both the sensory component of two-component systems and as the transducer of the signal to the response regulator altering gene expression (Parkinson 1993; Hoch and Silhavy 1995; Appleby *et al.* 1996; Raivio and Silhavy 2001a; West and Stock 2001). To allow this transduction of an external signal across the cell membrane to the response regulator, HK proteins are usually integral membrane proteins (Hoch and Silhavy 1995; Appleby *et al.* 1996; Raivio and Silhavy 2001a). They consist of an N-terminal sensory domain and a C-terminal kinase domain separated by the cellular membrane and joined together by a membrane spanning domain. It has also been shown that HKs exist as homodimers (Surette *et al.* 1996; Tomomori *et al.* 1999). This dimerization is known to be necessary for the kinase activity of the CheA HK (Surette *et al.* 1996). Transfer of a phosphate from CheA~P to ADP was dependant on the concentration of CheA~P, with transfer occurring rapidly at high CheA

concentrations, and more slowly at lower concentrations, where CheA~P would be forced to exist as a monomer.

The transduction of an environmental signal to the response regulator via the HK protein occurs through a conserved set of reactions catalyzed by specific domains, some of which exhibit sequence homology of varying degrees to each other within the protein. The first of these domains is the sensory domain, which is able to detect a specific environmental signal through an unknown mechanism (Parkinson 1993; Goudreau and Stock 1998; West and Stock 2001; Raivio and Silhavy 2001a). The observation that there is no sequence conservation amongst sensory domains underscores the fact that they are able to detect a wide variety of signals. These include such diverse stimuli as osmolarity, pH, the presence of various ions or metabolites, and phytochrome mediated photoreception. (Goudreau and Stock 1998; Raivio and Silhavy 2001a). The actual mechanism by which the detection of the stimuli occurs is currently not well understood, however two possibilities have been proposed. The first of these is direct sensing, in which the stimulatory molecule binds directly to the sensory domain of the HK. An example of this is the oxygen sensing FixL protein, which is able to bind oxygen directly by virtue of a haem group it contains (Raivio and Silhavy 2001a). This is opposed to indirect sensing, in which the stimulus is transduced to the HK by a distinct sensory molecule. This is illustrated by the NtrB HK, which responds to nitrogen limitation. In this system, NtrB is bound by the unmodified PII protein, signaling high nitrogen levels. The primary sensor of nitrogen levels modifies PII, allowing it to alter NtrB activity under activating conditions (Raivio and Silhavy 2001a). Also, HK proteins may respond to several distinct signals. The CpxAR two-component system is involved in the

responses to diverse stimuli. The stresses which can induce this system include envelope stress, adhesion to or contact with a biotic or abiotic surface, and entry into the stationary growth phase (DiGiuseppe and Silhavy 2003). However it is not clear at present how these different signals are detected. Two possibilities could be imagined, firstly that each signal is sensed differently, or alternativly that CpxA is able to detect a general stress produced by all three signals. In response to these signals, the Cpx system activates a diverse array of genes. Among these genes are the *pap* genes which are involved in the biogenesis of pili (Jones *et al.* 1997; Hung *et al.* 2001). The CpxAR system also controls the expression of many proteins involved in stress response, including the protease DegP, folding factors DsbA, PpiA/RotA, and PpiD, regulatory protein CpxP, and its own operon, *cpxAR* (Danese *et al.* 1995; Raivio and Silhavy 1999).

Subsequent to the detection of the activating signal, it must be transduced across the membrane to the HK catalytic domain. This is believed to be accomplished by the transmembrane domain, as mutations in this region interfere with the transduction of the signal and the activation of the catalytic domain (Raivio and Silhavy 2001a). Upon receipt of the signal, the catalytic domain becomes activated, resulting in a change in its enzymatic activity. The catalytic domain is the site of the autokinase activity of the HK, and when activated it is able to phosphorylate itself on a conserved histidine residue (Parkinson 1993; West and Stock 2001). This phosphate moiety is later transferred to the response regulator, allowing it to function as a transcriptional regulator.

Several other conserved domains are known to exist within HK proteins. These are the ATP-binding and kinase domains, which contain the F, N, G1 and G2 boxes, and the H-box, which is usually located at the N-terminal end of this domain (Kim and Forst

2001; Bilwes *et al.* 1999). As a whole the kinase domain is able to recognize and bind to ATP, allowing the transfer of a phosphate from this molecule to the conserved histidine residue, which is located in the H-box (Bilwes *et al.* 1999). Site directed deletion mutagenesis has allowed greater understanding of the functions of the F, N, G1 and G2 boxes, revealing that they are involved in both ATP binding, and the autokinase activity (Hirschman *et al.* 2001).

This identification of conserved domains amongst HKs allowed for their subdivision into five major types based on sequence homology (Kim and Forst 2001). Types I and II are similar and contain orthodox kinase domains. This means they contain all five of the previously mentioned domains. Type I can be further subdivided into types IA, IB, and IC based on less dramatic but still significant differences in homology. Both CpxA and BaeS are type IA HKs. In contrast to type I HKs, type III HKs and type IV HKs are missing the F-box, while the G2-box is shortened, and the N-box contains conserved and definitive changes to its consensus sequence. These are termed unorthodox kinases. Finally, type V HKs include the CheA-like chemosensors, which contain an orthodox kinase, however the H-box is located at the N-terminus of the protein and is separated from the kinase domain by an intervening sequence (Garzón and Parkinson 1996). Striking differences in the distribution of these HK types are noted amongst differing bacterial and archaeal species. While several bacterial species, including *E. coli* and *Bacillus subtilis* are known to contain all five types, type I is the predominant bacterial type. Type III is more prevalent in Gram positive bacteria, and types II and IV are less commonly found. Amongst the Archaea, it is noted that the HK complement is often dominated by a single type, such as type II in *Archaeoglobus fulgidus,* and type III

in *Methanobacterium thermoautotrophicum* (Kim and Forst 2001). This illustrates the origin of HK proteins, and their initial diversification, in bacteria, with isolated incidents of gene transfer leading to their presence in archaea with distinct distributions.

#### 1.1.1.2 Characteristics of Response Regulator (RR) Proteins.

Response regulators allow the bacterial cell to adapt to the stimuli detected by the HK by altering the transcription of a defined subset of genes, known as the system's regulon. This is achieved by an alteration in the activity of the RR as a transcriptional regulator in response to its phosphorylation (Parkinson 1993; Hoch and Silhavy 1995; Appleby *et al.* 1996; Raivio and Silhavy 2001a; West and Stock 2001). RRs were initially identified as a class of proteins when homology was observed between *Salmonella typhimurium* CheY and proteins involved in chemotaxis, membrane protein synthesis, and sporulation in *B. subtilis* and *E. coli* (Stock *et al.* 1985).

Structurally, RRs are usually cytoplasmic proteins, and consist of an amino-terminal (N-terminal) receiver domain and a carboxy-terminal (C-terminal) output domain (Parkinson 1993; Appleby *et al.* 1996; Raivio and Silhavy 2001a). As was described previously, phosphorylation of the RR occurs after phosphorylation of the conserved histidine residue in the HK. This phosphorylation reaction occurs at a conserved aspartate residue. Perhaps surprisingly, the catalytic activity for this reaction resides in the receiver domain of the RR itself, and not in the HK (Lukat *et al.* 1992). This was demonstrated by showing that the RRs CheY and CheB are able to autophosphorylate in the presence of small phospho-donors in the complete absence of their cognate HK, CheA. One of these small phospho-donors, acetyl phosphate, has been

suggested to serve as a global regulatory signal in *E. coli* (McCleary *et al.* 1993). This was suggested as the gene expression mediated by some two-component systems is not completely removed by the creation of null mutations in the HK. This observation has also been termed crosstalk, and is suggested to be a result of RRs being phosphorylated by non-cognate HKs. If acetyl phosphate is indeed serving as a global regulator, it is unclear what cellular response or pathways it is involved in (McCleary *et al.* 1993). Despite these suggestions, it has been shown by kinetic studies that CheY is able both to bind CheA~P more tightly than acetyl phosphate, and also that the phospho-transfer reaction is much more efficient in the presence of CheA~P than with small phospho-donors (Mayover *et al.* 1999). Other studies have analyzed the potential for crosstalk between specific two-component systems in *E. coli* (Verhamme *et al.* 2002). The two-component systems studied were UhpBA, PhoRB, NtrBC, and ArcBA, and no crosstalk was observed *in vivo,* when all of the systems were intact. It was found that deleting NtrB, the HK of the NtrBC system, resulted in some activation of NtrC by the other systems, but presumably this does not represent a natural situation.

Although it has long been believed that phosphorylation of the conserved aspartate residue in the receiver domain of a RR would lead to an alteration in its conformation, this was not demonstrated until fairly recently. Phosphorylation has been observed to result in the dimerization of the RR FixJ (Birck *et al.* 1999). It has also been observed that RR's tended to contain two topologically distinct domains, one at the C-terminus, and one at the N-terminus. Deletion of the N-terminal domain, known as the receiver domain and the site of phosphorylation, was found to result in the activation of the C-terminal domain, which is the effector domain responsible for DNA binding and

activation of the regulon (Grimsley *et al.* 1994). This suggested that the receiver domain was responsible for repressing the activity of the effector domain. Supporting this, and affirming our understanding of the phospho-relay archetype, phosphorylation of the receiver domain's conserved aspartate was observed to relieve this inhibition (Ireton *et al.* 1993). It has now been shown through the determination of a crystal structure that phosphorylation of both *B. subtilis* Spo0A and *E. coli* CheY results in a conformational alteration to those proteins (Lewis *et al.* 1999; Lee *et al.* 2001).Phosphorylation of Spo0A was found to alter the positioning of two residues, Thr84 and Phe103 which are located near the site of phosphorylation, and the hinge connecting the two domains (Lewis *et al.* 1999). This heavily implies that an alteration of the association between the receiver domain and the effector domain is responsible for the activation that occurs in response to phosphorylation.

RRs can be subdivided into different families based on sequence similarity. These are the OmpR, NtrC, and FixJ families, named after their most well characterized member (Hakenbeck and Stock 1996; Raivio and Silhavy 2001a). The OmpR family is the largest, and contains both the CpxR and BaeR RRs, which will be discussed in more detail in sections 1.2.2.2 and 1.3. This family is characterized by a receiver domain located at the N-terminus, which as discussed is believed to hold the protein in an inactive conformation prior to phosphorylation, and a 150 amino acid DNA-binding domain at the C-terminus (Hakenbeck and Stock 1996). The DNA-binding domain of this family is unique in structure, consisting of a winged-helix-turn-helix motif (Martínez-Hackert and Stock 1997). This is differentiated from the more common helix-turn-helix DNA binding motif by the presence of two loops, or wings, located to either side of the

recognition helix. These wings are also able to interact with DNA upon binding. In *E. coli*, OmpR-type RRs usually activate the expression of genes transcribed by the $\sigma^{70}$ RNA polymerase (RNAP) by making direct contact with the α subunit of the RNA polymerase (RNAP) core enzyme (Slauch *et al.* 1991; Hakenbeck and Stock 1996). Members of the OmpR family are also known to function as repressors depending on the precise location of their binding site in relation to the promoter. There is also evidence that the specificity of the OmpR protein is not due solely to its ability to recognize specific DNA sequences (Ohashi *et al.* 2004). Two of the two component systems in *E. coli* are responsive to osmolarity, EnvZ-OmpR which controls the porins OmpC and OmpF, and KdpDE, which regulates the potassium transporter KdpABC. It was found that OmpR was capable of recognizing the *kdpABC* promoter, but was unable to interact properly with RNAP to activate transcription. Mutations were identified in the OmpR RNAP interacting domain which were able to compensate for this and allow transcription, suggesting that spatial configuration of the promoter was also important for activation (Ohashi *et al.* 2004).

Of the three RR families, the FixJ family is the least well characterized (Raivio and Silhavy 2001a). It is similar in overall structure to the OmpR family, consisting of a C-terminal DNA-binding domain and a N-terminal receiver domain. However, the DNA binding domain is shorter than that of the OmpR family, consisting of only 100 amino acids, and there is no significant sequence similarity between them. It is also different in structure, containing a helix-turn-helix DNA-binding motif. This differs from the winged-helix-turn-helix, as it lacks the paired wings that characterize the latter motif. The determination of X-ray structures revealed that conformational alterations to the N-

terminal domain upon phosphorylation allow the protein to dimerize (Gouet *et al.* 1999; Birck *et al.* 1999). As FixJ crystals were unsuitable for structural determination, small-angle X-ray scattering has been employed to observe alterations in protein conformation on phosphorylation (Birck *et al.* 2002). It was found that phosphorylation led to conformational changes in an area of the protein near the helix-turn-helix (HTH) DNA binding motif, the α4-β5-α5 interface. This allowed homodimerization to occur, which is necessary for full activity (Birck *et al.* 1999). The α4-β5-α5 interface also contacts the linker region which connects the receiver and the effector domains, leading to the suggestion that this contact could result in the disruption of the interaction between the N and C-terminal domains, relieving the repression on DNA binding (Birck *et al.* 2002). This idea was later demonstrated by mutational studies showing that the α4 and α5 helices are involved in signal propagation between the two domains (Saito *et al.* 2003). As with the OmpR family, FixJ RRs control the transcription of promoters recognized by the $\sigma^{70}$ RNAP (Hakenbeck and Stock 1996). Although the mechanism by which these regulators alter gene expression is presently unknown, like OmpR, they are believed to function at the level of polymerase recruitment.

The NtrC family of RRs is characterized by the presence of three domains, a DNA-binding domain at the C-terminal, a N-terminal receiver domain, and an ATPase domain between them. As with the FixJ family, the DNA binding domain contains a helix-turn-helix domain, however unlike the other two families, it activates the transcription of genes controlled by promoters that are transcribed by $\sigma^{54}$ RNAP. The NtrC family is also unique in its mechanism of activation, in that it is involved in the transition from the closed to open transcription initiation complex (Hakenbeck and Stock

1996). The hydrolysis of ATP provides energy for this reaction, which is otherwise thermodynamically unfavourable (Wedel and Kustu, 1995). This function is dependent on the ATPase activity of the central domain, which is in turn dependant on the protein forming a tetramer. This occurs upon phosphorylation of the receiver domain by the HK, as unphosphorylated NtrC is a homodimer (Hakenbeck and Stock 1996). Oligomerization is believed to stabilize the proteins binding to DNA, a phenomenon termed cooperative binding. In contrast to the OmpR and FixJ classes of response regulator, which upon phosphorylation become activated as a result of altered contact between the N and C-terminal domains, NtrC RRs are activated by strengthened contact between the N-terminal domain and the central ATPase domain (Harrod *et al.* 2004). It is not currently known how the activity of the ATPase domain is able to facilitate the formation of an open complex (Raivio and Silhavy 2001a). Despite this however, it has been observed that all activators of $\sigma^{54}$ dependent promoters share the central ATPase domain, suggesting that open complex formation by $\sigma^{54}$ RNAP is thermodynamically unfavourable.

## 1.2 Envelope Stress Response

### 1.2.1 Characteristics of the Gram Negative Envelope.

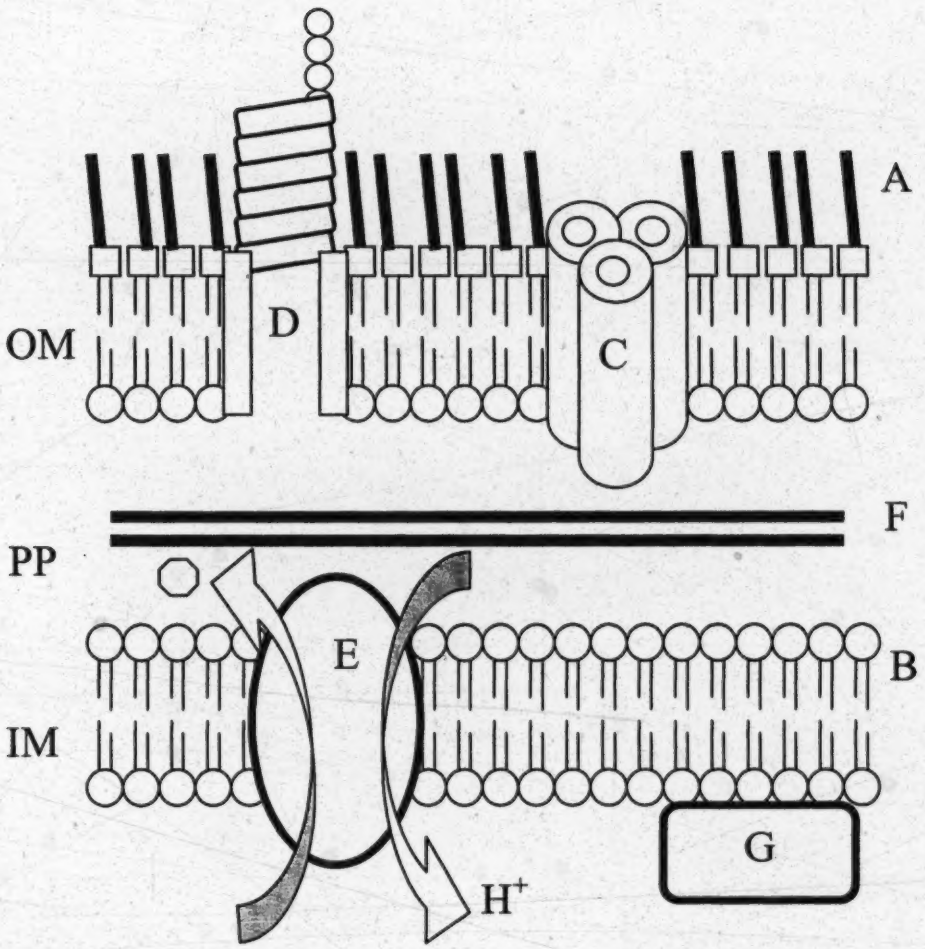
The envelope is the outermost component of the Gram-negative bacterial cell, and as such it serves as the point of contact between the cell and the environment. The membranes of the envelope have a number of specialized functions which include transport and cell motility. The envelope is particularly exposed to environmental stress,

and the cell must maintain its integrity if it is to survive. This is the purpose of the envelope stress response.

The Gram-negative envelope is composed of two lipid bilayers, the inner and outer membranes, which enclose the periplasmic space (Figure 1.2). As its name implies, the outer membrane is the outermost part of the envelope. This membrane is composed of two types of lipid, lipopolysaccharides (LPS), which make up the outer lipid layer and phospholipids, which compose the inner leaflet (Nikaido 1996a). The phospholipids present in this membrane are similar to that of the inner membrane and consist of, in decreasing abundance, phosphatidylethanolamine, phosphatidylglycerol, and cardiolipin. LPS is unique to the outer leaflet of the outer membrane and is structurally distinct from phospholipids, consisting of the O-antigen polysaccharide which projects into the extra cellular medium and is joined to lipid A by the core oligosaccharide. The study of mutants deficient in LPS production has shown that this molecule is involved in the function of the membrane as a physical barrier, as LPS mutants display greater sensitivity to a variety of compounds, including hydrophobic antibiotics and detergents (Nikaido 1996a). The expression of LPS also acts as a virulence factor in pathogenic strains, as it is involved in evading the host immune response and assisting with adherence to epithitial cells. Upon cell death and subsequent lysis, the release of lipid A also acts as an endotoxin eliciting a strong immune response in humans. The outer membrane is also the location of an assortment of proteins which are involved in regulating the passage of chemicals into and out of the cell. In *E. coli*, these include the porins OmpC and OmpF. These porins do not have tight substrate specificity, and facilitate the passage of a variety

**Figure 1.2. The Gram-negative envelope.**

The leaflets of the outer membrane are distinct, with the outer leaflet consisting of O antigen linked to LPS (A) while the inner membrane is a phospholipid bilayer (B). Both the inner and outer membranes are characterized by a distinct set of integral membrane proteins such as porins (C) and pili (D) in the outer membrane, and antiporters (E) in the inner membrane. These membranes enclose the periplasm, which contains a layer of peptidoglycan (F). Proteins may associate with either of the membranes without penetrating the lipid bilayer (G).


A
OM
D
C
F
PP
E
B
IM
G
H+

of molecules into the cell. Another important pore located in the outer membrane is TolC, which is required for the export of antibiotics by many multidrug efflux pumps (Fralick 1996). Other important outer membrane localized proteins include those involved in cellular motility, such as the flagella, and those involved in cellular adhesion to a surface, such as pili. Motility is achieved by the ability of the long extracellular flagella to rotate, propelling the cell towards a chemical attractant, or in a non-directed manner (Harshey and Toguchi, 1996). Pili are also implicated in pathogenesis, as a particular pilus, the P-pilus has been found to be expressed predominantly in strains causing infections of the urinary tract (Kallenius *et al.* 1981).

The inner or cytoplasmic membrane is distinct from the outer membrane in both composition and function. In terms of composition, the lipid constituents of this membrane differ from those of the outer membrane (Raetz and Dowhan 1990). This is most noticeable in the absence of LPS in the inner membrane, however the phospholipid composition differs slightly as well with the outer membrane being slightly enriched in phosphatidylethanolamine (Nikaido 1996a). The lipid composition of the membrane is known to have a profound effect on the functioning of membrane associated proteins, as cells deficient for production of a particular lipid display characteristic problems with their proteins (Kadner 1996). For example, the efficiency of lactose transport by the permease LacY is decreased by 10 to 20 fold in a strain unable to produce the phospholipid phosphatidylethanolamine.

The inner membrane contains a substantial protein component, making up a full 50% of the of the membrane surface. Proteins associated with the inner membrane play important roles in numerous functions such as transport and cell motility. Since the

Gram-negative cell has two distinct membranes, in order to transport material either into or out of the cell it is necessary to cross both membranes. This is usually achieved by separate proteins in each membrane, as is the case in multidrug efflux. TolC is needed for the efflux of antibiotics. To accomplish this however, it associates with a specific transporter protein situated in the inner membrane (Koronakis 2003). Cell motility via chemotaxis is also influenced by inner membrane proteins, including the HK CheY. The phosphorylation state of CheY determines whether it is able to associate with the FliM switch protein. This in turn determines the direction of flagellar rotation (Kadner 1996). There is also a pH gradient across the inner membrane, and several classes of protein exploit this. These include of transport proteins, such as symporters and antiporters, and ATP synthase.

The inner and outer membranes enclose the periplasmic space, which is a viscous proteinaceous environment (Oliver 1996). Due to a high concentration of protein, diffusion within the periplasm is lower than that found in either an aqueous medium, or the cytoplasm (Brass *et al.* 1986). This space is characterized by the presence of a peptidoglycan layer, which is associated with the outer membrane. The major class of proteins present in the periplasm are the periplasmic binding proteins, which function with transporter proteins in the outer membrane and serve to concentrate their substrates for transport across the inner membrane. These include MalE and FecB which bind maltose and iron-citrate respectively. Other proteins in the periplasm are involved in diverse functions, including the catabolism of macromolecules, stress response and detoxification, which includes DegP, and the biogenesis of the envelope (Oliver 1996). In addition to this, numerous inner membrane proteins have a domain located in the

periplasm, such as HK proteins. The periplasm is also notable for being an energy-depleted environment, lacking a source of ATP. As a result of this, processes in the envelope requiring energy rely on the proton motive force present across the inner membrane.

### 1.2.2 General Characteristics of the Envelope Stress Response.

Envelope stress is regulated in *E. coli* separately from more general stress responses, and is controlled by an ExtraCytoplasmic function (ECF) σ factor, and a two-component system, CpxAR (Raivio and Silhavy 2000; Raivio and Silhavy 2001b). These systems control the expression of genes which encode proteases, protein folding factors, an alternate σ factor and genes associated with biosynthesis of membrane constituents, as well as genes, *rpoE* and *cpxAR,* allowing them to carry out autoregulation (Danese and Silhavy 1997; Connolly *et al.* 1997; Missiakas and Raina 1998). The expression of these regulons is aimed at alleviating the stress that the cells have been challenged with.

#### 1.2.2.1 The $\sigma^E$ Envelope Stress Response.

Sigma factors (σ) are proteins which are able to interact with the RNAP core enzyme and dictate its promoter recognition capabilities. Different σ factors will direct RNAP to different subsets of promoters. This serves as a useful method for changing the overall transcriptional capabilities of the cell. The ECF σ factors were identified as a separate class based on the observation of sequence homology between the *Streptomyces coelicolor sigE* gene and σ factor genes in a number of other organisms, including $\sigma^E$ in *E. coli* (Lonetto *et al.* 1994). This sequence homology was also reflected in functional

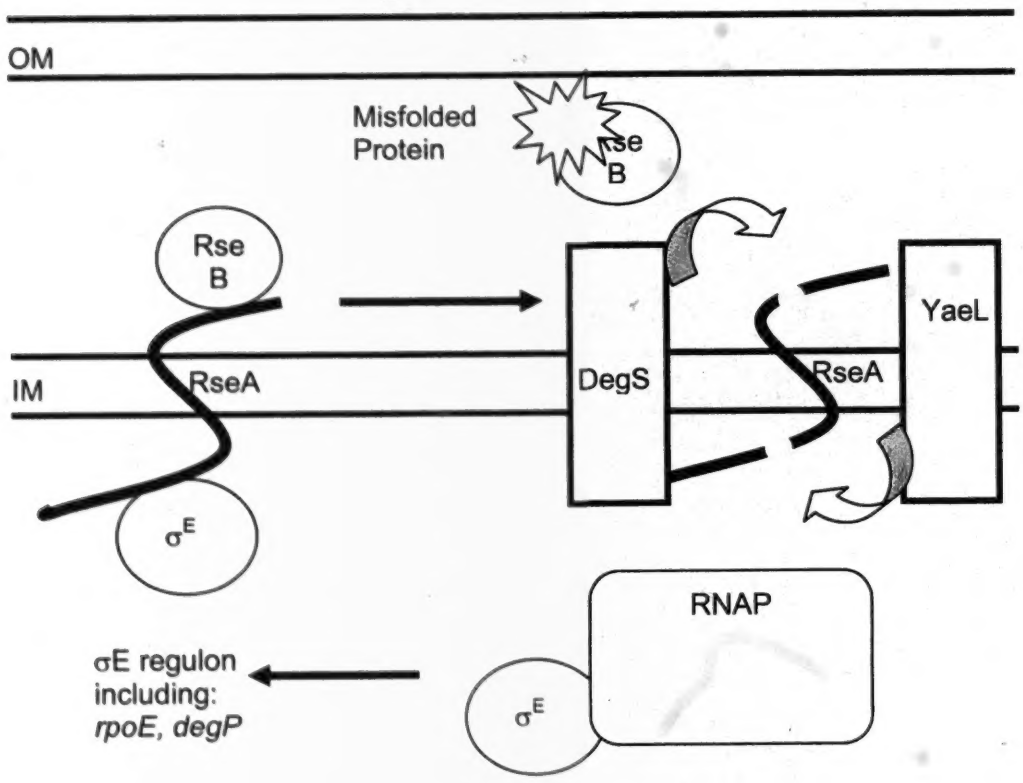
homology as many of these genes were known to encode σ factors involved in extracytoplasmic responses.

In *E. coli,* $\sigma^E$ was first implicated in stress response due to its regulation of *degP*. DegP was known to be a protease whose presence was required for survival at high temperatures (Strauch and Beckwith 1988; Lipinska *et al.* 1990; Kolmar *et al.* 1996). However, DegP was found to be regulated independently from $\sigma^H$ which was known to mediate the heat shock response in *E. coli* (Lipinska *et al.* 1988). This suggested that *degP* was controlled by a second, unidentified, heat shock response system. This new heat shock response system was later shown to be controlled by $\sigma^E$, which was shown to control the *degP* gene (Erickson and Gross 1989). To further understand the $\sigma^E$ response, searches were carried out for mutations which affected $\sigma^E$ activity. The results of these searches showed that $\sigma^E$ responded to changes in the level of outer membrane proteins (OMPs), as well as to mutations that resulted in the misfolding of OMPs (Mecsas *et al.* 1993). This led to the suggestion that $\sigma^E$ was involved in monitoring the state of proteins in the outer membrane, and was therefore an envelope stress response.

The mechanism by which $\sigma^E$ monitored OMPs was not determined until the regulation of the $\sigma^E$ response was described in more detail. $\sigma^E$ is a cytoplasmic protein, however it was observed that it was negatively regulated by the integral inner membrane protein RseA, and to a lesser extent by the periplasmic protein RseB (De Las Peñas *et al.* 1997). The model that emerged from this study was that under non-inducing conditions, $\sigma^E$ would be sequestered at the inner membrane by RseA (Figure 1.3). This interaction has been shown to involve the 4.2 region of $\sigma^E$ which is responsible for binding the -35 region during promoter recognition (Tam *et al.* 2002). This interaction prevents $\sigma^E$ from

**Figure 1.3. The $\sigma^E$ envelope stress response.**

Under non-inducing conditions (left), $\sigma^E$ is sequestered at the inner membrane (IM) by a complex of RseA and RseB. Misfolded proteins (starburst) titrate RseB away, leading to sequential cleavage of RseA by DegS and YaeL, which results in the destabilization of the RseA-$\sigma^E$ complex and the release of $\sigma^E$ into the cytoplasm, where it interacts with the RNAP holoenzyme. This leads to the transcription of the $\sigma^E$ regulon.


OM
Misfolded Protein
Rse B
Rse B
IM
RseA
DegS
RseA
YaeL
σ^E
RNAP
σE regulon including: rpoE, degP
σ^E

associating with the RNAP core to form the RNAP holoenzyme, preventing the transcription of the $\sigma^E$ regulon (Collinet *et al.* 2000). The association of $\sigma^E$ and RseA is stabilized by RseB binding to the periplasmic portion of RseA, resulting in RseA:RseB having 2-3 fold greater affinity for $\sigma^E$ than RseA alone. Under inducing conditions, RseB is able to associate with denatured proteins in the periplasm, titrating it away from interaction with RseA (Collinet *et al.* 2000). Subsequent to this, RseA is cleaved first in its periplasmic domain by the inner membrane associated protease DegS and subsequently in its membrane spanning or cytoplasmic domain by YaeL (Ades *et al.* 1999, Alba *et al.* 2002). This results in the destabilization of the RseA/$\sigma^E$ complex, resulting in the release of $\sigma^E$ into the cytoplasm. Once released from its association with RseA, $\sigma^E$ interacts with RNAP, directing the holoenzyme to the promoters it recognizes, thereby activating expression of the regulon (De Las Peñas *et al.* 1997).

The $\sigma^E$ regulon contains a number of genes whose expression alleviates envelope stress. This includes the protease DegP, which is able to degrade misfolded periplasmic proteins. Also induced is the periplasmic peptidyl prolyl isomerase, FkpA, which is involved in protein folding in the envelope, and additional stress response genes such as *skp, dsbC,* and *surA* (Missiakas and Raina 1998; Raivio and Silhavy 2001b, Dartigalongue *et al.* 2001). SigmaE also regulates the expression of $\sigma^H$ leading to an upregulation of its regulon as well. The $\sigma^H$ regulon includes such protein folding factors as DnaK/J and GroEL/S and protein degradation factors like Lon and ClpS (Gross 1996). It has been observed that the regulon of $\sigma^E$ contains members which, while not appearing to be directly involved in stress response, are involved in envelope biogenesis, such as *htrM, lpxD* and *ecfA,* all of which are part of the LPS synthesis pathway (Dartigalongue

*et al.* 2001). Several genes of unknown function are also known to be $\sigma^E$ regulon members, including *ecfB, ecfD, ecfF, ecfG, ecfH, ecfI, ecfJ, ecfK, and ecfL,* all of which are predicted to localize to the envelope. To date, the remaining members of the known $\sigma^E$ regulon are regulatory or autoregulatory in nature. These are the $\sigma^E$ regulators *rseABC*, and the $\sigma^E$ gene itself, *rpoE*.

#### 1.2.2.2 The Cpx Envelope Stress Response.

The CpxAR system is a two-component system which controls a second envelope stress response in *E. coli*. The Cpx response overlaps to some extent with that of $\sigma^E$, however they are distinct systems which respond to different inducing cues with divergent regulons (Raivio and Silhavy 1999). The initial identification of the Cpx system was made in a mutational screen to identify genes involved in the elaboration of the F-pilus, which is involved in conjugative plasmid transfer (McEwen and Silverman 1979). *cpxA* was identified in this screen, with the mnemonic standing for conjugative plasmid expression. The original mutants generated were unable to elaborate the F-pilus, and were therefore deficient in conjugation and the attachment of certain bacteriophage. However, instead of only being involved in the regulation of pilus assembly, later studies noted that mutations to the *cpx* genes also altered the protein compositions of both the inner and outer membranes (McEwen and Silverman 1982; McEwen *et al.* 1983). The specific effect observed was a deficiency in the levels of murein lipoprotein and the OmpF porin. Analysis of the nucleotide sequences of *cpxA* and *cpxR* revealed them to be a HK and RR respectively (Weber and Silverman 1988; Dong *et al.* 1993). Since the identification of CpxAR as a two-component system suggested in broad strokes how gene

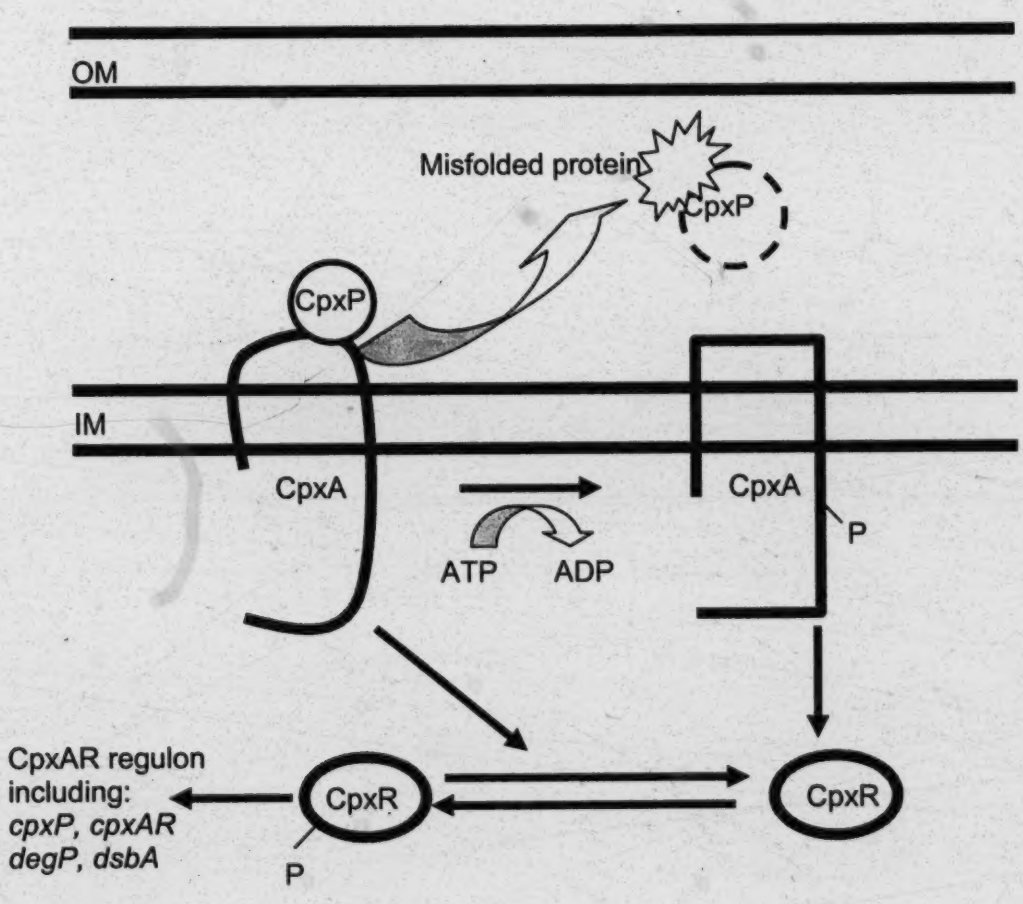
activation would be achieved, it became important to further characterize both the stimuli this system was responding to, and the identities and functions of the members of its regulon.

An important observation for elucidating the function of the Cpx system was the discovery that mutations in genes encoding both HK CpxA and RR CpxR affected the transcription of *degP* (Danese *et al.* 1995). Specifically, it was found that a gain-of-function mutation in *cpxA* increased transcription of *degP*, while null mutations in both *cpxA* and *cpxR* decreased its transcription. Since DegP is a periplasmic protease responsible for degrading misfolded proteins, this suggested that the CpxAR system was involved in envelope stress response (Figure 1.4). Further evidence that CpxAR was a second envelope stress response was provided by a study in which the presence of a permanently activated form of CpxA, CpxA*, was found to alleviate some known envelope stresses (Cosma *et al.* 1995). Specifically, CpxAR suppressed toxicity resulting from the overexpression of misfolded proteins in the periplasm in a DegP dependant manner.

The CpxAR envelope stress response pathway is known to respond to a wide variety of environmental stimuli. These include adhesion to a surface, changes in pH, alterations in membrane composition, and the overexpression of NlpE, an outer membrane lipoprotein (Otto and Silhavy 2001, Danese and Silhavy 1998, Mileykovskaya and Dowhan 1997, Snyder *et al.* 1995). Further studies suggested that the Cpx system responds to misfolded proteins associated with the inner membrane, such as misfolded P-pilus subunits, whereas $\sigma^E$ responds to protein misfolding associated with the outer membrane (Raivio and Silhavy 1999). The CpxAR dependent expression of folding factors has been found to

**Figure 1.4. The CpxAR envelope stress response.**

Under non-inducing conditions (left), the phosphatase activity of CpxA maintains CpxR in a non-phosphorylated state. CpxA is activated by the presence of misfolded proteins (starburst) in the periplasm. This is thought to relieve repression by the periplasmic CpxP protein, possibly by titrating it away from an association with CpxA. This activates CpxA as an autokinase, leading to its phosphorylation on a conserved histidine residue. This phosphate (P) is transferred to the RR CpxR activating it as a transcriptional regulator. This results in the transcription of a regulon that includes periplasmic protein folding factors.


OM
Misfolded protein
CpxP
CpxP
IM
CpxA
CpxA
P
ATP
ADP
CpxAR regulon
including:
cpxP, cpxAR
degP, dsbA
CpxR
P
CpxR

be necessary for full elaboration of the P-pilus, and therefore, a pathogenicity factor (Hung *et al.* 2001). Furthermore, CpxR positively regulates the *pap* locus itself, which encodes the p-pilus.

As has been described, the Cpx regulon includes the DegP protease, the expression of which is able to alleviate toxicity due to accumulation of misfolded proteins in the periplasm. Other genes regulated by CpxAR include the disulfide bond forming protein DsbA, and the peptidyl prolyl isomerases PpiA and PpiD, all of which are involved in the folding of envelope proteins (Raivio and Silhavy 1999). CpxAR is also involved in the expression of virulence genes in some pathogenic strains of *E. coli,* such as uropathogenic *E. coli* (Hung *et al.* 2001). A study using the consensus sequence for promoter recognition by CpxR~P has identified many new potentially CpxAR regulated genes (DeWulf *et al.* 2002). This study suggested that the total number of genes regulated by CpxAR may be over 100, implying that the Cpx response may be rather more diverse than had been previously suggested.

The CpxAR system is also notable for containing a third component, the periplasmic protein CpxP, belying the term two-component system. CpxP was initially identified via a transpositional screen using the λplacMu53 bacteriophage (Danese and Silhavy, 1998). This protein was found to be able to repress expression of the CpxAR regulon in a manner that was dependent on the CpxA sensing domain, leading to the suggestion of an interaction between the two proteins (Raivio *et al.* 1999). This suggestion was strengthened by the observation that tethering CpxP to the inner membrane prevented induction of the CpxAR pathway during spheroplasting, a physical disruption of the outer membrane which results in the depletion of the constituents of the

periplasm (Raivio *et al.* 2000). Recently it has been observed that upon induction of the CpxAR system, CpxP is degraded in a manner dependent on DegP, suggesting a role for this protease in maintaining activation of the CpxAR system after induction (Buelow and Raivio, 2005). Interestingly, the only known homologue of CpxP is the Spy protein (section 1.3.2.1), another periplasmic protein of unknown function which is known to be jointly regulated by the CpxAR system and the BaeSR system (Raivio *et al.* 2000, Raffa and Raivio, 2002).

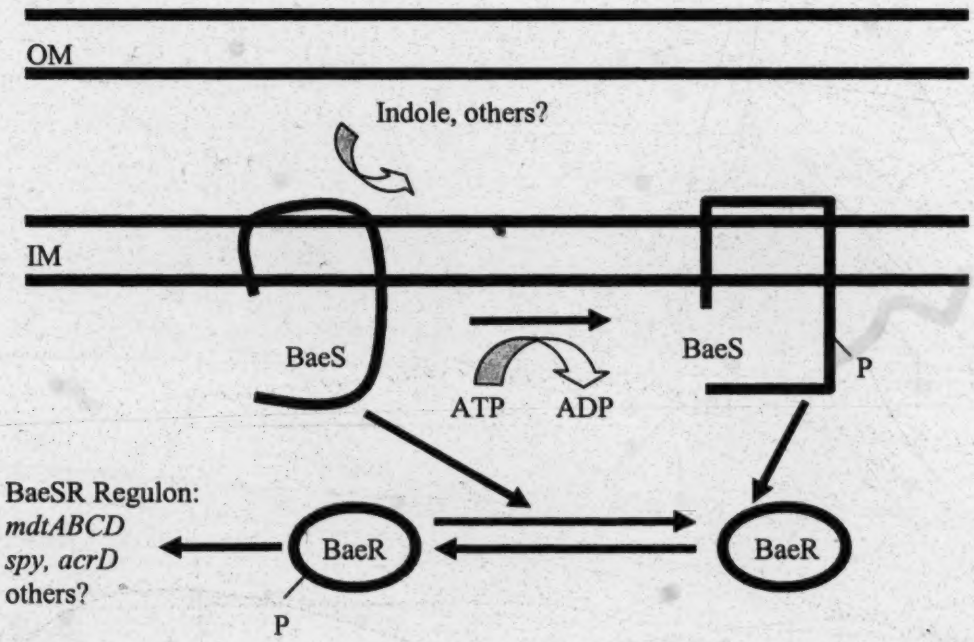
## 1.3 The BaeSR Two-Component System

### 1.3.1 The initial identification of BaeSR.

The *E. coli* proteins BaeS and BaeR comprise a recently identified two-component system that is currently poorly defined both in terms of its inducing cues and the cellular response it mediates (Figure 1.5). The BaeSR system was first identified by screening random genomic fragments for the ability to suppress mutations in the EnvZ and PhoR/CreC HKs (Nagasawa *et al.* 1993). Since EnvZ regulates porin expression in response to osmolarity levels, and PhoR responds to phosphate levels, this complementation should not necessarily be seen as implying a functional homology in terms of perceived stress, but rather homology of the kinase activity of the HK protein. It could be postulated that in the absence of their cognate sensors, OmpR and PhoB may be able to phosphorylate themselves using phosphate moieties on other, non-cognate, HKs. However, 'crosstalk' such as this between members of different two-component regulatory systems has been shown not to play a significant role in wild type cells

**Figure 1.5. The BaeSR envelope stress response.**

HK BaeS is activated by the presence of the envelope stressor indole. This activates it as an autokinase, leading to its phosphorylation on a conserved histidine residue. This phosphate is transferred to the RR BaeR activating it as a transcriptional regulator. The regulon of this system is not well characterized, but it is known to include the multidrug efflux pump encoding genes *mdtABCD*, as well as *spy* and *acrD*.


OM
Indole, others?
IM
BaeS
ATP
ADP
BaeS
P
BaeSR Regulon:
*mdtABCD*
*spy, acrD*
others?
BaeR
P
BaeR

(Verhamme *et* al. 2002). It was observed to some extent in cells with mutations in their HKs, lending credence to this suggestion.

After this initial identification, other studies independently identified the BaeSR system and implicated it as being involved in different responses. One of these experiments utilized transposon mutagenesis to identify factors involved in the regulation of the *spy* gene in response to the envelope stress of spheroplasting (Raffa and Raivio 2002). Another study identified BaeSR in a screen of genomic fragments for the ability to increase novobiocin resistance in a Δ*acrAB* background (Baranova and Nikaido 2002).

### 1.3.2 The BaeSR Regulon

#### 1.3.2.1 BaeSR is involved in envelope stress.

As stated above, the BaeSR system was first implicated as being involved in the envelope stress response by its regulation of the Spy protein. Spy is a small periplasmic protein of unknown function. It was first identified by its differential expression in spheroplasts, where it is induced, and intact cells, where it is undetectable (Hagenmaier *et al.* 1997). This observation suggested that Spy was a stress protein. Since the CpxAR system has also been shown to be induced by spheroplasting, this and the upregulation of *spy* in a strain containing the constitutively active *cpxA24* mutation suggested that the expression of Spy may be part of the CpxAR response (Raivio *et al.* 2000). This was shown to be the case by examining the activity of a *spy-lacZ* reporter in a variety of Cpx backgrounds. Furthermore, *spy* induction by spheroplasting was found to be dependent on CpxAR (Raivio *et al.* 2000). The only known protein to which Spy shows any homology

is CpxP (section 1.2.2.2), a negative regulator of the CpxAR system (Danese and Silhavy 1997; Raivio *et al.* 2000). Even in this case however, the homology observed is only limited, with a 26.5% amino acid identity, and a 42.8% similarity. Despite this however, Spy and CpxP do not appear to play the same role in the regulation of CpxAR activity, as Cpx reporters are not affected by overexpression of Spy (Raivio *et al.* 2000). Furthermore, the induction of Spy in response to envelope stress was only partially dependent on CpxAR, and was independent of the $\sigma^E$ pathway. This suggested the existence of a third envelope stress response system.

In order to locate this putative third envelope stress response, transpositional mutagenesis was carried out on a strain containing a *spy:lacZ* reporter (Raffa and Raivio 2002). This revealed that an insertion in the BaeS gene was able to substantially increase expression of the reporter. This ability was shown to be independent of CpxR activity. In addition to this, Spy expression was shown to be induced by several other envelope stresses, such as the overexpression of the p-pilus subunit, PapG, and the presence of indole. Induction in response to these stresses by BaeR and CpxR overlapped to some extent, with the response to PapG being mediated primarily by CpxR, and the response to indole depending mainly on the presence of BaeR (Raffa and Raivio 2002). Furthermore, induction of *spy* by BaeR did not occur in response to overexpression of NlpE, a powerful activator of the CpxAR system. This demonstrates that although the BaeSR and CpxAR pathways overlap to some degree in terms of both their regulons, and their inducing cues, they are in fact distinct envelope stress response systems. Supporting the suggestion that BaeSR is involved in envelope stress is the observation that strains with mutations in *baeR* are more sensitive to envelope stressors than either wildtype cells, or

strains mutant only for *spy* (Raffa and Raivio 2002). This second observation suggests that other genes involved in the envelope stress response may be controlled by BaeSR. For this reason more work is needed in order to further characterize both the BaeSR regulon, and the cues it is activated in response to.

#### 1.3.2.2 BaeSR is involved in Multidrug Resistance.

Other studies have also implicated BaeSR as being involved in cellular responses distinct from those of the CpxAR pathway. Screening of random genomic fragments for the ability to increase resistance to novobiocin in a Δ*acrAB* background revealed *baeR* (Baranova and Nikaido 2002). Resistance to the bile salt deoxycholate was also increased in this strain. This increase in resistance was shown to be due to the *mdtABCD* genes, which are part of an operon that contains *baeSR* as well. Consistent with the suggestion that BaeR regulates the expression of *mdtABCD* was the observation that BaeR binds to the *mdtA* promoter (Baranova and Nikaido 2002).

The *mdtABCD* genes encode a multidrug efflux pump, wherein MdtA is a membrane fusion protein (MFP), MdtB/MdtC are resistance-nodulation-division (RND)-type transporters, and MdtD is a major facilitator (MF) type transporter (Nishino and Yamaguchi 2001b). Research using different fragments of the *mdtABCD/baeSR* operon has shown that MdtD is not required to confer resistance to novobiocin and deoxycholate (Nagakubo *et al.* 2002). Furthermore, overexpression of MdtAC was shown to confer resistance to bile salts alone, while a plasmid carrying *mdtAB* did not increase resistance at all. In order to get an increase in resistance to both bile salts and novobiocin, both MdtB and MdtC were required in conjunction with MdtA. This observation led to the

conclusion that the transporter of the Mdt efflux pump is actually a heteromultimer of MdtB and MdtC (Nagakubo *et al.* 2002). This conclusion makes the MdtBC transporter unique amongst RND-type exporters, which are usually homomultimers.

Other studies of two-component system control of drug resistance have confirmed that the BaeSR system is able to confer resistance to a variety of different antibiotics. These include a wide range of β-lactams such as oxacillin and cloxacillin (Hirakawa *et al.* 2003a). BaeR mediated resistance to these compounds was found to be dependent on both the *mdtABC* genes, and on *acrD*, a second multidrug exporter. Similar results were found when this same group looked at resistance to a wider variety of compounds. They found that BaeR overexpression led to resistance to novobiocin, SDS, and deoxycholate (Hirakawa *et al.* 2003b). Once again this resistance was dependant upon both MdtABC and AcrD. Transcript analysis was used to demonstrate that BaeR overexpression led to an increase in the amount of both *mdtA* and *acrD* mRNA, indicating transcriptional control.

Recently it has been noted that the control of *mdtABCD* and *acrD* in response to indole is coordinately regulated by BaeSR and CpxAR (Hirakawa *et al.* 2005). Deletion of *baeSR* was found to reduce dramatically the indole inducible phenotype of these genes, while deletion of *cpxAR* was found to reduce it to a lesser extent. Indole induction was eliminated by the deletion of both systems. Thus it was found that the primary regulation in response to indole was BaeR dependant, while CpxR was able to play a modulating effect, enhancing expression of *mdtA* and *acrD*. This was supported by the observation that both BaeR and CpxR were able to bind the promoter regions of these genes. DNaseI protection assays were utilized to determine the BaeR and CpxR binding sites in these

promoters, and it was found that they were in close proximity to each other and the transcription start site (Hirakawa *et al.* 2005). BaeR was found to bind to a single site closer to the start codon, whereas CpxR bound several sites immediately upstream of this, as would be expected of a primary regulator, and a protein modulating its effect.

#### 1.3.2.3 Other possible functions of the BaeSR pathway

As it is known to be involved in the regulation of two multidrug efflux pumps, and responds to envelope stressors such as indole and spheroplasting, the BaeSR pathway has been identified as an envelope stress response system. There have been several studies of interest regarding the function BaeSR system, but they do not fit comfortably into the above sections. The first is a phenotype microarray analysis of a strain harbouring a deletion in BaeSR. This study looked at differences in growth between wildtype cells and a *baeSR* deletion strain over a wide variety of conditions. It was found that this strain had increased sensitivity to myricetin, as well as gallic acid, nickel chloride, and sodium tungstate (Zhou *et al.* 2003). This supports the idea that BaeSR is involved in both multidrug efflux, and stress response.

Another interesting study suggests that indole may be used as an extracellular signal by *E. coli* (Wang *et al.* 2001). It was observed that *E. coli*-conditioned medium was able to activate a variety of genes including *astD*, *tnaB*, and *gabT* which are involved in amino acid metabolism. The activating signal for this response was purified and identified as indole. As has been previously described, indole is known to be an activating signal for the BaeSR pathway. Could BaeSR therefore be what is regulating

these genes in response to indole, and is this occurring in a cell density-dependent manner?

#### 1.3.2.4 Other putative regulon members

In addition to established regulon members *spy*, *mdtABCD* and *acrD,* several other genes that lack a clear connection to envelope stress or multidrug resistance have been suggested to belong to the BaeSR regulon. One study screened a genomic library for fragments that increased the expression of a *lacZ* reporter gene in response to BaeR overexpression, resulting in the identification of the promoter regions of the *yicO* and *ygcL* genes (Baranova and Nikaido 2002). Neither of these genes has been characterized in terms of function or operon status, however *yicO* has homology to the nucleobase:cation symporter family, specifically, the xanthine/uracil permeases. Likewise, *ygcL* has homology to phophatidylinositol kinases and DNA dependent protein kinases.

Recently, microarray analysis has been utilized to look at the genome wide effects of BaeR overexpression (Nishino *et al.* 2005). While the expression of 59 genes appeared to be increased by overexpression, only seven of these were both decreased by BaeR deletion, and induced by indole. Of these seven, the only novel regulon member was *ycaC,* which is speculated to be a hydrolase based on sequence similarity, and is homologous to pyrazinamidase and carbamoylsarcosine amidohydrolase. A comparison of the promoter regions of the known regulon members was carried out and a putative consensus binding sequence was proposed. Only 50% identity was observed to this consensus sequence in the region of the *ycaC* promoter speculated to be the BaeR binding

site, while the well established regulon members such as *spy* and *mdtA* had binding sequences which exhibited between 94% and 100% sequence identity. It is notable that homology searches did not reveal any sequences matching the consensus sequence in any of the other 59 genes initially identified. Thus although several genes, *yicO*, *ygcL* and *ycaC,* recently suggested as belonging to the BaeSR pathway are of undetermined function, those which have been well established appear to respond to envelope stress, such as *spy*, or be multidrug efflux pumps, such as *mdtABCD,* and *acrD.*

## 1.4 Multidrug Resistance

### 1.4.1 General Characteristics of Efflux Pumps.

Among the hazards that may confront a bacterial cell are antibiotics and other toxic compounds. These are chemicals which are able to either kill bacteria, usually through the inhibition of vital cellular processes, or interfere with their division and growth. Antibiotics are present in the natural environment, being secreted into the soil by some bacteria such as *B. subtilis*, possibly in order to discourage the growth of competing species. With the advent of modern medicine however, antibiotics have truly become a common challenge faced by a wide variety of bacteria. There has long been concern that overuse of antibiotics could lead to the emergence and spread of pathogenic strains with resistance to one or more of the antibiotics commonly used to treat infections. Unfortunately, this process is already occurring, as studies have reported dramatic increases in the rates of resistance observed in many bacterial infections from 1995 to 2000 (Shannon and French 2004). This trend has dire ramifications for the future of

medical treatment for a large variety of infections, and as such the mechanism of bacterial antibiotic resistance is an important topic for research.

One of the most common means of acquiring antibiotic resistance to an assortment of antibiotics is the expression of multidrug efflux pumps (Levy 1992; Levy 2002). These systems consist of an outer membrane pore, such as TolC, a periplasmic membrane fusion protein, and an inner membrane transporter, which is responsible for the specificity of the complex, and is what they are classified by. Far from being specific pumps which export only one class of molecule, multidrug efflux pumps have a broad substrate specificity, and are able to export substances which bear little structural resemblance to each other. These substrates can include diverse, though primarily hydrophobic and amphipathic, antibiotics, biocides and organic solvents, as well as bile salts (Levy 1992; Levy 2002). Although this observation calls into question what the substrates of these pumps would be in the natural environment, they have in common the fact that they are hazardous to the cell. Their expulsion therefore serves a defensive purpose, leading to a generalized resistance (Okusu and Nikaido 1996; Thanassi *et al.* 1997). Despite this broad range of substrates however, different multidrug efflux pumps do have different ranges of substrates. Perhaps as a consequence of this, most bacterial species encode several different multidrug efflux systems. *E. coli* for example contains some 285 transporter proteins, 6-18% of which are estimated to be involved in drug resistance (Paulsen *et al.* 1998). This large number of drug transporters can be subdivided into the resistance-nodulation-division (RND), major facilitator (MF), ATP binding cassette (ABC), and small multidrug resistance (SMR) families based on energy source, homology, and substrate range of their inner membrane component (Nikaido 1996,

Tulkens *et al.* 2000). Of the above families, only the ABC pumps are primary active transporters; the other classes function as either symporters or antiporters to transport their substrates. The RND, MF and SMR families are classified by their structure, with the SMR class being the smallest, consisting of only four transmembrane domains. The RND and MF pumps are larger, consisting of twelve or fourteen transmembrane regions, and are distinguished by the large periplasmic domains typical of RND pumps (Tulkens *et al.* 2000).

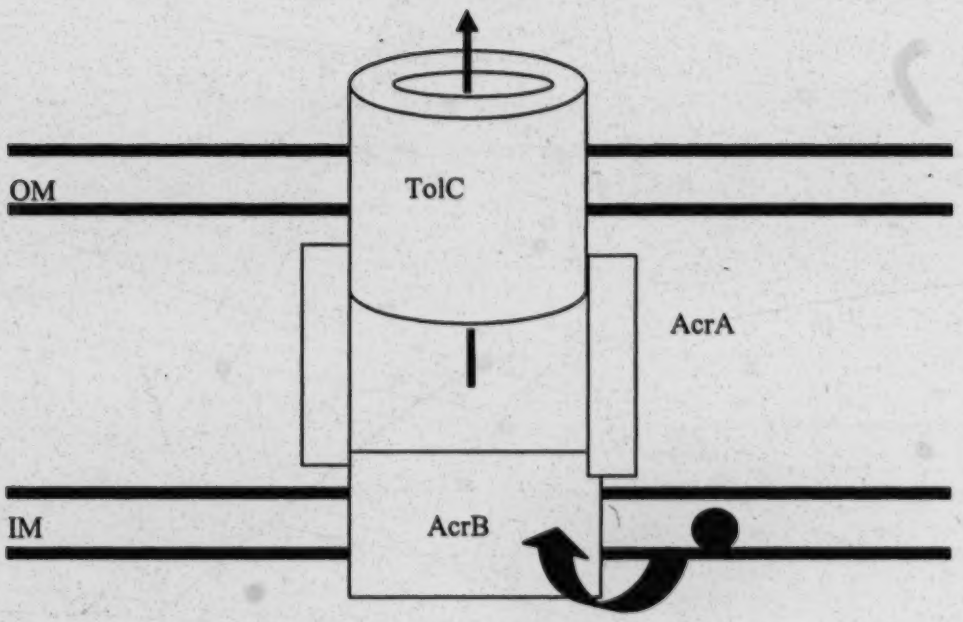
Drug export in Gram-negative organisms is complicated by the need to export the toxic compounds across two separate membranes, and the periplasmic space between them (Thanassi *et al.* 1995; Zgurskaya and Nikaido 2000). Failure to do this would result in a build-up of toxins in the periplasm, from where they would be able to quickly pass back into the cytoplasm. To achieve export across both membranes, multidrug export systems consist of a number of different proteins acting together.

#### 1.4.1.1 Structure and Function of Multidrug Efflux Pumps.

In general, Gram-negative multidrug efflux pumps consist of three proteins associated to form a functional complex. These proteins are an inner membrane transporter, an outer membrane pore, and a periplasmic membrane fusion protein (Okusa and Nikaido 1996). One of the major efflux systems of *E. coli* is the AcrAB multidrug efflux pump, and as one of the best characterized systems, it is often considered the archetype of multidrug efflux pump structure and function for RND type transporters (Figure 1.6). This system consists of the AcrB transporter and the AcrA membrane fusion

**Figure 1.6. The AcrAB multidrug efflux pump.**

This system is typical of multidrug efflux pumps in *E. coli*. The RND transporter AcrB is able to remove hydrophobic molecules from within the periplasm, the inner membrane and the cytosol, transporting them though an inner channel. These molecules are transported out of the cell by the outer membrane porin TolC. AcrA is believed to facilitate a close interaction between TolC and AcrB, allowing transport to bypass the periplasm. This diagram is intended to illustrate the gross topology of the AcrAB multidrug efflux pump and not to represent its structure or conformation.


OM
TolC
AcrA
IM
AcrB

protein, associated with the TolC outer membrane pore (Ma *et al.* 1995; Fralick 1996). This system confers resistance to a wide variety of compounds, including the antibiotics novobiocin and tetracycline, and compounds likely to be encountered in the intestinal habitat of *E. coli*, such as bile salts and fatty acids (Rosenberg *et al.* 2003).

The inner membrane transporter of the AcrAB system is AcrB. The inner membrane transporter is the only component that differs substantially between the RND and MF families (Okusa and Nikaido 1996). AcrB belongs to the RND family, which has a large periplasmic domain lacked by MF transporters (Van Bambeke *et al.* 2000). The action of this transporter is powered at the expense of proton motive force; it functions as a proton antiporter, importing protons as molecules in its substrate range are exported. This generalization has been confirmed by the observation that AcrB is only able to function in the presence of a pH gradient across the inner membrane, which *in vivo* is present across the inner membrane (Zgurskaya and Nikaido 1999). Structurally, AcrB functions as a homo-trimer, and consists of three domains, the transmembrane domain, the periplasmic pore, and the TolC docking domains. The transmembrane domain consists of 12 membrane spanning α-helices, which associate in a loose bundle. These helices are believed to contain a proton translocation pathway, ferrying hydrogen ions from the periplasm to the cytoplasm and providing the energy for substrate export (Murakami *et al.* 2002). On the exterior of the bundle there is a groove leading from the cytoplasm to the pore domain, representing a possible pathway for substrate translocation. Located in the periplasm, the pore domain encloses a large central cavity (Murakami *et al.* 2002). Situated at the interface between the outer leaflet of the inner membrane and the periplasm are vestibules which lead to the central cavity. These

vestibules represent a second possible route for substrate translocation. Additionally the pore domain is believed to be the site of association with AcrA, and a deep cleft on its exterior is believed to be the location of this interaction. The TolC docking domain has been suggested to facilitate interaction with the TolC protein, which is located in the outer membrane. Structurally, the TolC docking domain adopts a funnel like conformation, the diameter of which corresponds closely to that of the bottom of TolC leading to suggestions these proteins may interact (Murakami *et al.* 2002).

There has been some uncertainty about the mechanism by which transport across the membrane is achieved. It has been proposed that the transporter takes its hydrophobic substrates from within the membrane itself, rather than exporting them from the cytoplasm directly (Van Bambeke *et al.* 2000). This model suggests that the drug enters the central chamber of the transporter from the membrane via the vestibules, where it is then expelled out of the cell. Recent studies using AcrD have shown that some multidrug efflux pumps are in fact able to capture substrates from the cytoplasm, as well as from the periplasm and the inner membrane (Aires and Nikaido, 2005). Furthermore, it was noted that some substrates, such as streptomycin, were captured from one location, but not the other, suggesting that the uptake capacity of AcrD is distinct in these two compartments. Also unclear has been the ability of AcrB to act on structurally diverse substrates including compounds which are cationic, anionic, zwitterionic, lipophilic, and amphiphilic (Nikaido, 1996). This question has been clarified by the publication of the crystal structure of AcrB in complex with various substrates (Yu *et al.* 2003). It was found that the interior of the AcrB central cavity is a structurally diverse region, and that distinct substrates are positioned differently such that they are stabilized by hydrophobic

interactions with the chamber wall, as well as electrostatic interactions with charged residues for charged molecules.

AcrA is the membrane fusion protein (MFP) of the AcrAB system. MFP's were shown to be a protein family through both sequence homology and conserved function, with initial observations revealing they were involved in efflux, and were associated with both an inner and outer membrane partner (Dinh *et al.* 1994). Structurally, AcrA is known to be a very elongated protein (Nogales *et al.* 2001). Experiments to determine the function of AcrA have shown that it is sufficient to cause an exchange of lipids between two separate vesicles, suggesting that AcrA is able to cause an association between two membranes, or perhaps even limited fusion (Zgurskaya and Nikaido, 1999). The publication of the crystal structures of AcrB and TolC have called into question the need for this however, as these two proteins are large enough to span the periplasm on their own (Murakami *et al.* 2002, Koronakis *et al.* 2000) AcrA is able to interact with AcrB and TolC, however the complex between AcrB and TolC may be unstable in the absence of AcrA (Elkins and Nikaido 2003, Touzé *et al.* 2004). This observation suggests that MPFs are involved in creating a closer association between the inner and outer membrane components of multidrug efflux pumps. This would be explain why they are essential for transport, and potentially explain the observation that the efflux of antibiotics normally bypasses the periplasm entirely (Thanassi *et al.* 1995).

The final component of a multidrug efflux pump is the outer membrane pore. The outer membrane pore is required for transport across the outer membrane. This achieves the expulsion of the pump substrate from the cell, bypassing the periplasm completely and thus preventing a potentially toxic buildup of drugs in this space. The pore which

associates with the AcrAB efflux pump is the TolC pore which was found to be required for the proper functioning of this pump (Fralick 1996). TolC was shown not to affect the level of AcrAB expression, and so it is not involved in the regulation of these proteins. This suggests that TolC is part of the efflux complex itself. In addition to AcrAB, TolC also serves as the outer membrane component of the MF type transporter EmrBA (Nikaido 1996). This result makes it clear that TolC is able to interact with several different efflux pumps, serving as a generalized outer membrane transport protein. Structurally, TolC consists of three domains, a β-barrel domain, an α-helical domain, and an equatorial α-β domain located in the periplasm (Koronakis 2003). The β-barrel domain is positioned in the OM, and is not obstructed, making the interior of the TolC protein continuous with the extra-cellular solution. The α-helical domain extends below this domain into the periplasm. This domain is an α-helical barrel, which encloses a very large interior space (Hughes *et al.* 2000). Several of the helices of this domain taper inwards near its end so as to seal the interior of the porin from the periplasm, preventing it from being continuous with the exterior of the cell. Upon association with an inner membrane partner such as AcrB these helices are thought to shift in position relative to one another, opening the pore (Anderson *et al* 2002.). Lastly, the equatorial α-β domain is located in a band encircling the α-helical domain and may be the site of contact with AcrA. Taken as a whole, the extended barrel-like structure of TolC makes it easy to envision the mechanism by which export across the outer membrane is achieved.

### 1.4.1.2 Regulation of Multidrug efflux

As a result of the importance of efflux systems for the survival of the cell in hazardous conditions, and as they are also known to excrete some metabolic intermediates such as adenylosuccinate, citrate, and isocitrate, control of their expression is of particular importance (Helling *et al.* 2002). Perhaps as a result of this, multidrug efflux pumps are controlled by a number of different systems, in response to a variety of signals. Again this can be illustrated by examining the AcrAB system. AcrAB expression is controlled by several regulatory systems at both a global and a local level. Globally, *acrAB* production is controlled by the activators MarA, SoxS, and Rob (White *et al.* 1997). Activation by MarA occurs in response to the presence of salicylate, while SoxS activates *acrAB* in the presence of super oxide free radicals (Cohen *et al.* 1993; Ding *et al.* 1996). Activation by Rob occurs in response to its binding of bile salts and fatty acids, substrates of the AcrAB pump (Rosenberg *et al.* 2003). Acting at the local level is the repressor AcrR (Ma *et al.* 1996). Some inducers of AcrAB, such as general stress, also induce AcrR expression to an even greater extent. This induction of a repressor of *acrAB* transcription may prevent overexpression of the AcrAB efflux pump, which may result in toxicity due to a loss of metabolites. Additional evidence for the toxic effects of AcrAB overexpression is provided by the difficulty experienced by researchers attempting to clone *acrAB* into a high copy number plasmid, and by the observation that *E. coli* strains which up regulate *acrAB* experience slower growth rates than wild-type strains (Ma *et al.* 1996).

In addition to these levels of control, AcrAB expression has been shown to be controlled by the LuxR homologue SdiA (Wei *et al.* 2001). This raises the possibility that

efflux can be controlled by quorum sensing, and may illuminate the induction of AcrAB observed to occur in stationary phase. Regulation similar to this is also seen for other multidrug efflux pumps in *E. coli.* This includes the EmrAB efflux pump, which is involved in resistance to the antibiotics thiolactomycin and nalidixic acid. Expression of EmrAB is repressed by the EmrR protein, which binds to the *emrRAB* promoter, preventing transcription (Lomovskaya *et al.* 1995; Xiong *et al.* 2000). This binding is alleviated by the presence of drugs which induce expression of the EmrAB pump.

In addition to these modes of regulation, many multidrug efflux pumps are controlled by two-component systems providing a mechanism of tying efflux pump expression to environmental conditions. This includes the *E. coli* efflux pumps EmrKY and YhiUV, whose expression is controlled by the EvgAS two-component regulatory system (Nishino and Yamaguchi 2001a; Nishino and Yamaguchi 2002). It has been demonstrated that overexpression of the EvgA RR leads to an increase in the expression of the YhiUV pump. This results in an increase in the minimum inhibitory concentrations of several antibiotics, including deoxycholate, doxorubicin, and erythromycin. Despite this however, the environmental signals sensed by the EvgS HK protein remain undetermined. The increase in expression of multidrug efflux pumps fostered by overexpression of two-component systems has been shown to result in resistance to a variety of antibiotics (Hirakawa, *et al.* 2003). Amongst the two-component systems included in this study were CpxAR and BaeSR. In the absence of the AcrAB efflux pump, CpxAR overexpression was found to increase resistance to kanamycin, novobiocin, and the bile salt deoxycholate while BaeSR overexpression conferred a higher level of resistance to novobiocin, as well as sodium dodecyl sulfate (SDS) and

deoxycholate. For the BaeSR system, these resistance phenotypes were abolished in the presence of BaeSR overexpression in strains that contained deletions in the multidrug efflux pumps *mdtABC* and *acrD* (Hirakawa, *et al.* 2003). Although the environmental signals to which these systems respond is not well characterized, these observations demonstrate the ability of bacteria to respond to environmental signals with the expression of multidrug efflux genes, increasing their resistance to various antibiotics.

Multidrug efflux is also known to be induced by toxic compounds and pump substrates. An example of this is the Bmr efflux pump of *B. subtilis*. The expression of this pump is controlled by the activity of its regulatory protein BmrR (Ahmed *et al.* 1994). The ability of BmrR to bind the *bmr* promoter was shown to be affected by the presence of rhodamine and TPP, substrates of the pump. This provides an example of a multidrug efflux pump being specifically upregulated by the presence of its substrates.

The chemical indole is a byproduct of *E. coli* metabolism, and at higher concentrations is thought to disrupt the bacterial envelope. Recently, this compound has been found to induce the expression of a variety of efflux genes, including *acrD, mdtA, acrE, emrK,* and *yceL* (Hirakawa *et al.* 2005). Induction of *acrD* and *mdtA* via indole was found to depend on BaeSR and CpxAR, with the Cpx system serving to modulate the primary regulation exerted by the Bae system. It is not know if the BaeSR system is able to regulate other genes involved in the response to chemical stressors.

## 1.5 Proposed line of Research

### 1.5.1 A screen to identify BaeSR regulated genes.

Currently, very little is known about either the inducing signals for, or the regulon controlled by, the BaeSR two-component system. It is known to respond to envelope stressors such as indole and PapG overexpression in a manner that is distinct from the CpxAR and $\sigma^E$ pathways (Raffa and Raivio 2002). However, its established regulon includes only *spy*, the multidrug efflux pump *mdtABCD*, and the MF transporter *acrD* (Raffa and Raivio 2002; Hirakawa *et al.* 2003b). It has been shown that the BaeSR system must control additional genes involved in envelope stress as a *bae* mutant shows sensitivity to envelope stressors beyond that of the wildtype strain, or a *spy* mutant (Raffa and Raivio 2002). Furthermore, it is difficult to conceive why the expression of multidrug exporters MdtABCD and AcrD, which export hydrophobic compounds from within the inner membrane, would be induced by signals like the expression of misfolded periplasmic proteins like PapG. This suggests that more work should be done both in identifying the BaeSR regulon, and in understanding the role of this response in the physiology of *E. coli*.

To further the understanding of the BaeSR pathway, this thesis intended to identify additional members of the BaeSR regulon. We initiated a project screening for genes affected by overexpression of BaeR prior to the publication of a microarray study looking at the same stimulus (Nishino *et al.* 2005). To this end, I carried out a genetic screen using transpositional mutagenesis to interrupt random genes on the *E. coli* chromosome with a λp*lac*Mu53 element which bears a promoterless *lacZ* gene. This was done by utilizing the λp*lac*Mu53 phage, which utilizes the phage Mu transposition functions to insert non-specifically into the chromosome (Bremer *et al.* 1985). The insertion of this transposon into a gene can result in the production of an operon fusion

wherein a *lacZ* reporter gene is controlled by the promoter of the interrupted gene. This creates a convenient method to examine the induction of the target gene via β-galactosidase assays. Strains selected for further study showed differential levels of β-galactosidase activity in the presence and absence of BaeR overexpression. This was expected to reflect the activation, or potentially the repression, of the gene interrupted by the *lacZ* reporter construct by BaeSR.

In order to verify that this was actually the case, and to ensure that only one insertion was present in the chromosome, P1*vir* transduction was utilized to transfer the insertion, selected for by kanamycin resistance, into an *E. coli* MC4100 background. The resulting strains were retransformed with the BaeR overexpression plasmid, which resulted in the return of the arabinose mediated phenotypic difference. If this was observed, the potentially BaeSR regulated genes were further characterized by sequencing outwards from the λp*lac*Mu53 transposon, in order to determine their identities. Once a number of genes that exhibited phenotypes consistent with potential BaeSR regulation were identified, electrophoretic mobility shift assays (EMSAs) were carried out in an attempt to confirm that these were indeed regulon members by demonstrating BaeR binding to their promoters.

In order to further characterize the newly identified putative BaeSR regulon members, expression studies were carried out via two techniques. Firstly, β-galactosidase assays were carried out using the λp*lac*Mu53 insertion strain in the absence of BaeR overexpression, and the same λp*lac*Mu53 insertion in the presence of *baeS1::Tn10cam*, a *baeS* allele containing an insertion which results in a gain of function mutation conferring increased activity of the BaeSR pathway. Secondly, these λp*lac*Mu53 insertion strains

were examined by β-galactosidase assays under both non-inducing conditions and inducing conditions, which for this project were 2mM indole, and the recently identified inducing cues 10μg/mL myricetin and 10mg/mL sodium tungstate. This series of experiments was intended to confirm BaeSR regulation of the newly identified genes, and to examine the effectiveness of differing inducing cues.

### 1.5.2 Concluding Remarks

Although the BaeSR system has been implicated in both the envelope stress response and in multidrug resistance, a comprehensive model for the induction and physiological function of this pathway has yet to emerge. In order to address this gap, my research was intended to identify further genes controlled by BaeSR. By increasing our knowledge of the BaeSR regulon this project was intended to give insight into the role the BaeSR response plays in the physiology of *E. coli*, and perhaps into the environmental cues that would induce it.

# 2. Materials and Methods

## 2.1 Bacterial Strains, Phage, Plasmids, and Growth Conditions

### 2.1.1 *E. coli* strains

The *E. coli* strains used and created in the course of this study are listed in table 2.1.

### 2.1.2 Growth and maintenance of *E.coli* strains

*E. coli* liquid cultures were grown in LB medium (10g Bacto-Tryptone(BD), 5g Bacto-yeast extract(BD), 5g NaCl(EM Sciences) brought to 1L in dH2O), or MacConkey broth (35g/L Difco) at 37°C in an orbital shaker at 225rpm. Solid cultures were grown on LB medium (as above with 1.5% agar w/v), or MacConkey Agar (40g/L Difco) at 37°C. When antibiotic selection was required, the following final concentrations were used, kanamycin (Kan)(100μg/mL), ampicillin (Amp)(100μg/mL), and chloramphenicol (Cam)(25μg/mL). Frozen stocks of *E. coli* strains were prepared by mixing equal amounts of an overnight liquid culture of the strain in question with 10% glycerol in $dH_2O$, resulting in a final glycerol concentration of 5% and subsequently freezing at -80°C.

### 2.1.3 Phage used in this study

The phage used in this study are listed in table 2.2.

**Table 2.1 *Escherichia coli* strains**

| ***E. coli* strain** | **Genotype** | **Reference or Source** |
|---|---|---|
| MC4100 | Wild type | Casadaban (1976) |
| TR51 | MC4100 *cpxR::spc* | Raivio *et al.* (1999) |
| TR61 | TR50 ara$^R$ Tn*10cam* | Raivio |
| TR776 | MC4100 *baeS1::*Tn*10cam* | Raivio *et al.* (2002) |
| TR886 | MC4100 λRS88 [*spy-lacZ*] *baeR1::Kn* | Raivio *et al.* (2002) |
| AC31 | TR530 (pMAL-c2:*baeR*) | Chan (unpublished) |
| MC4100Δ*baeR* | MC4100 with *baeR* deleted | Hirakawa (2005) |
| MC4100Δ*baeSR* | MC4100 with *baeSR* deleted | Hirakawa (2005) |
| CA22 | MC4100 ara$^R$ Tn*10cam* (pCA14) | Andrews (unpublished) |
| CO2 | MC4100 *yggC*::λp*lac*Mu53 | This Study |
| CO4 | MC4100 *mdtA*::λp*lac*Mu53 | This Study |

| *E. coli* strain | Genotype | Reference or Source |
|---|---|---|
| CO5 | CO4 (pCA14) | This Study |
| CO6 | CO2 (pCA14) | This Study |
| CO8 | MC4100 *yliA*::λp*lac*Mu53 | This Study |
| CO10 | CO8 (pCA14) | This Study |
| CO16 | MC4100 *yjgX* intergenic region::λp*lac*Mu53 | This Study |
| CO17 | CO16 (pCA14) | This Study |
| CO22 | MC4100 $ara^R$ Tn*10cam* | This Study |
| CO23 | CO22 *fecD*::λp*lac*Mu53 | This Study |
| CO24 | CO23 (pCA14) | This Study |
| CO25 | CO22 *yjgX* intergenic region::λp*lac*Mu53 | This Study |
| CO26 | CO25 (pCA14) | This Study |
| CO37 | TR884 (pCA14) | This Study |
| CO46 | CO17 $ara^R$ Tn*10cam* | This Study |
| CO47 | CO6 $ara^R$ Tn*10cam* | This Study |

| *E. coli* strain | Genotype | Reference or Source |
|---|---|---|
| CO56 | Δ*baeSR cpxR::spc* | This Study |
| CO57 | MC4100 *baeS1::*Tn*10cam* | This Study |
| CO58 | TR776 *yggC*::λp*lac*Mu53 | This Study |
| CO59 | TR776 *yliA*::λp*lac*Mu53 | This Study |
| CO61 | TR776 intergenic region::λp*lac*Mu53 | This Study |
| CO63 | TR776 *mdtA*::λp*lac*Mu53 | This Study |
| CO67 | CO5 ara$^{R}$ Tn*10cam* | This Study |
| CO75 | TR776 *fecD*::λp*lac*Mu53 | This Study |
| CO88 | MC4100Δ*baeR fecD*::λp*lac*Mu53 | This Study |
| CO89 | MC4100Δ*baeR yjgX* intergenic region::λp*lac*Mu53 | This Study |
| CO90 | MC4100Δ*baeR mdtA*::λp*lac*Mu53 | This Study |
| CO91 | MC4100Δ*baeR yggC*::λp*lac*Mu53 | This Study |
| CO92 | MC4100Δ*baeR yliA*::λp*lac*Mu53 | This Study |

**Table 2.2 Phage strains**

| Phage strain | Description | Reference or Source |
|---|---|---|
| λ p*lac*Mu53 | Phage Mu based transposable element carrying a promoterless *lacZ* gene and a $Kan^R$ determinant | Weinstock (1985) |
| λ pMu507 | *c*Its875 *Sam*7 $MuA^+$ $MuB^+$ Helper phage for increasing the efficiency of λ p*lac*Mu53 infection | Allet (1977) |
| P1*vir* | Generalized transducing phage | |

### 2.1.4 Phage preparation and quantification

#### 2.1.4.1 λ Phage plaque assay

A 5mL overnight liquid culture (O/N) of *E. coli* MC4100 was prepared. This was harvested by centrifuging for 10 minutes at room temperature (RT) and resuspended in 2.5mL 10mM $MgSO_4$(BDH). Ten-fold dilutions of λ phage were prepared in λ dilution buffer (1.21g Tris Base[Roche], 1.2g $MgSO_4$, 0.1 gelatin[Knox], brought to 1L in $dH_2O$ and sterilized by autoclaving). A volume of 0.1mL of cells and 0.1mL of phage (λ p*lac*Mu53, Figure 3.1, or λ pMu507, various dilutions) were combined in plating tubes. The tubes were gently mixed, then allowed to sit for 10 minutes at RT. Three mL pre-melted (55°C) top agar (as for solid culture medium above, with 0.7 w/v agar) was added to the tubes, which were shaken gently, then poured onto pre-warmed (37°C) LB plates and incubated O/N at 37°C. The next day, the number of plaques on each plate was counted to determine the titre of the initial lysate.

#### 2.1.4.2 High titre λ liquid lysates

A 5mL O/N culture of *E. coli* MC4100 was harvested by centrifugation and resuspended in 2.5mL 10mM $MgSO_4$. Fifty μL aliquots were placed into each of 5 plating tubes. To these were added either 1, 2, 3, 4, or 0 individual plaques gathered using a sterile Pasteur pipette. The tubes were vortexed briefly, then allowed to sit for 5 minutes

at RT. Two milliliters of LB containing 10mM $MgSO_4$ were added, and the tubes were incubated shaking at 37°C until lysis occurred, or for 4 hours. After this, 0.1mL chloroform was added and the tubes were vortexed for 30 seconds. The resulting lysates were transferred to sterile centrifuge tubes and were centrifuged at 4500g for 10 minutes. The supernatants were removed to sterile screw capped tubes, and stored at 4°C. Plaque assays were performed to determine the titres as described above.

### 2.1.5 Plasmids and cloning vectors used and created in this study

All of the plasmids and cloning vectors used and created for this study are listed in Table 2.3. The BaeR overexpression plasmid pCA14 was created by Chris Andrews, by subcloning *baeR* from pCA3 into pBAD18 using EcoRI. pCA3 had been created by cloning the 800bp fragment amplified by PCR using primers BaeR5'Bam and Bae3'Pst and containing the *baeR* coding sequence into pCR2.1TOPO. This vector takes advantage of overhanging Adenosine residues left by Taq polymerase on the ends of the PCR fragment to improve cloning efficency. pMAL-c2:*baeR* was made by Anita Chan also utilizing the primers BaeR5'Bam and Bae3'Pst to amplify a 800bp PCR product consisting of the BaeR coding sequence. This fragment was cloned into pMAL-c2 using BamHI and PstI.

**Table 2.3 Plasmids and Cloning vectors**

| Plasmid | Description | Marker | Reference/Source |
|---|---|---|---|
| pBAD18 | Cloning vector with arabinose inducible *araBAD* Promoter | Amp | Guzman (1995) |
| pCA3 | pCR 2.1 TOPO derivative with *baeR* cloned downstream of P*lac* | Amp/Kn | Andrews and Raivio, unpublished |
| pCA14 | pBAD18 derivative with *baeR* cloned downstream of P*araBAD* | Amp | Andrews and Raivio, unpublished |
| pMAL-c2:*baeR* | pMAL-c2 derivative encoding MBP-BaeR fusion protein | Amp | Chan and Raivio, unpublished |

## 2.2 Transduction and transformation

### 2.2.1 Preparation of *E. coli* competent cells

#### 2.2.1.1 Electroporation competent cells

A 5mL O/N culture of the desired strain was prepared and subcultured 1:500 in 25mL of LB with the appropriate antibiotics. This was grown to an O.D. of 0.6 measured at 600nm, and harvested by centrifugation for 10 minutes at 4°C and 4024xg. The cell were resuspended in a 1X volume of cold $dH_2O$, and again harvested by centrifugation at 4°C and 4024xg, followed by resuspension in a 1/2X volume of cold $dH_2O$. This process was repeated, resuspending cells first in a 1/100X volume of cold $dH_2O$, then in a 1/500X volume of cold 10% glycerol. On the last step the cells were transferred to a microfuge tube, and were either used immediately or stored at -80°C.

#### 2.2.1.2 Chemically competent cells

A 5mL O/N culture of the desired strain was prepared and subcultured 1:50 in 5mL of LB with the appropriate antibiotics. This was grown to an $OD_{600}$ of 0.2 then harvested by centrifugation for 10 minutes at RT and 11688xg. The cells were resuspended in 1mL cold 'Magic Formula' (10mL 1M $CaCl_2$ (BDH), 80mL $dH_2O$, 10mL 1M 3-(N-morpholino) propanesulfonic acid (MOPS) (EM Science) pH6.5). This was incubated on ice for 30 minutes, then again pelleted by centrifugation at RT and 11688xg. The cells were gently resuspended in 0.2mL 'Magic formula' and transferred to a 1.5mL

microfuge tube. Cells prepared in this manner were either used immediately or stored at -80°C.

### 2.2.2 Transformation of *E. coli* cells

#### 2.2.2.2 Electroporation

Previously prepared frozen electro-competent cells (section 2.2.1.1) were thawed on ice. One hundred µL of these cells were combined with 1µL of purified plasmid DNA in a chilled electroporation cuvette (Molecular Bioproducts, 1mM). Electroporation was performed in a Biorad pulser on setting EC2. The cells were immediately transferred into 1mL of LB without antibiotics and recovered for 1 hour at 37°C in an orbital incubator shaking at 225 rpm. The cells were then pelleted by centrifugation at RT and 2700xg, the supernatant was decanted and the residual pellet was vortexed and plated on media with appropriate selection.

#### 2.2.2.2 Heat shock

Frozen competent cells prepared as described in section 2.2.1.2 were thawed on ice. One to five µL of purified plasmid DNA (section 2.3.1) were added and the mixture was incubated on ice for 10 minutes. The tube was then transferred to a water bath at 42.5°C for 30 seconds. Immediately after this, 1mL of LB was added and the cells were incubated for 1 hour at 37°C in an orbital incubator shaking at 225rpm. The cells were pelleted by centrifugation at RT and 2700xg, the supernatant was removed, and the residual pellet was vortexed and plated on media with appropriate selection.

### 2.2.3 P1*vir* Transduction

#### 2.2.3.1 Preparation of P1*vir* lysates

P1*vir* lysates were prepared as described previously (Silhavy *et al.* 1984). Briefly, a 5mL O/N culture of the donor strain was prepared with appropriate selection. This was subcultured 1:50 in 5mL of LB containing 0.2% glucose and 5mM $CaCl_2$, and incubated for 30 minutes at 37°C in an orbital incubator shaking at 225rpm. After this, 100μL of a P1*vir* lysate grown (via this procedure) on MC4100 was added. The culture was returned to the incubator for 2-3 hours at 37°C until lysis was observed. If complete lysis was not observed after 3 hours the next step was carried out. One hundred μL of chloroform were added to the tube which was vortexed for 30 seconds. The culture was pelleted by centrifugation at RT and 2700xg for 10 minutes and the supernatant was carefully transferred to a sterile screw-capped tube. The lysate was stored for future use at 4°C.

#### 2.2.3.2 P1*vir* Transductions

P1*vir* transductions were performed as described in Experiments with Gene fusions (Silhavy *et al.* 1984). Briefly, a 5mL O/N culture of the recipient strain was prepared with appropriate selection. The culture was harvested by centrifugation at RT and 2700xg for 10 minutes and resuspended in 2.5mL 10mM $MgSO_4$ with 5mM $CaCl_2$. In a small tube 100μL of the recipient cells and 100μL of a P1*vir* lysate grown on the donor strain were combined. This was incubated for 30 minutes at 37°C without shaking, after which 1mL LB containing 10mM citrate was added and the tubes were transferred

to a 37°C incubator for an additional 30 minutes. The cells were then harvested by centrifugation at RT and 2700xg for 10 minutes and plated on selective media.

## 2.3 DNA isolation, analysis and purification

### 2.3.1 Isolation of plasmid DNA from *E. coli*

Plasmids harboured in *E. coli* strains were isolated using the Sigma® GenElute™ Plasmid Miniprep Kit, following the manufacturer's instructions.

### 2.3.2 Polymerase chain reaction

All of the oligonucleotide primers utilized for this study are found in Table 2.4. The polymerase chain reaction was performed in 100μL volumes containing a final concentration of 1.5mM $MgCl_2$, 1X reaction buffer (Invitrogen), 100pM of each oligonucleotide primer, and 1μL of Taq polymerase (kindly provided by M.A. Pickard, University of Alberta). Template DNA was in the form of 1μL of purified plasmid, or a single colony of a bacterial strain, boiled in the reaction mixture for 5 minutes to lyse the cells.

Cycling was carried out in a Genius thermocycler (Techne). A typical reaction program was as follows: 30 cycles consisting of a 45 second 95°C denaturation stage, a 45 second 55°C annealing stage, and a 1 minute 72°C elongation stage. After the final cycle had been completed an extended elongation phase of 4 minutes was carried out at 72°C. For longer fragments, the extension time was lengthened. The annealing temperature was adjusted to suit individual primers.

**Table 2.4 Oligonucleotide Primers**

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| AcrD-F-ER1 | -CCGGAATTCTGCCTCCTACTGACCA AAGAA- | *acrD* promoter region (2585454-5472) | PCR fragment for EMSAs, DNaseI protection assay | This Study |
| AcrD-R-BH1 | -CGCGGATCCAATGGGGCGATCAAT AAAGA- | *acrD* promoter region (2585644-625) | PCR fragment for EMSAs | This Study |
| ARB1 | -GGCCACGCGTCGACTAGTACNNNN NNNNNNGATAT- | Varies, repetitive genome elements | Sequencing λp*lac*Mu53 insertions | Caetano-Anolles 1993 |
| ARB2 | -GGCCACGCGTCGACTAGTAC- | Varies repetitive genome elements | Sequencing λp*lac*Mu53 insertions | Caetano-Anolles 1993 |

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| BaeR5'Bam | -GGGATCCCCGAGTTACCAATCGAC GAAAAC- | BaeR coding region (2162302-2162322) | Cloning of *baeR* | Andrews and Raivio Unpublished |
| Bae3'Pst | -AACTGCAGTTTGCTGACCATGCGTGG- | BaeR coding region (2163102-2163082) | Cloning of *baeR* | Andrews and Raivio Unpublished |
| cmtBFbam | -CGGGATCCCGGTTTTAAGCGCCCCA CAT- | *cmtB* promoter region (3077119-137) | PCR fragment for EMSAs | This Study |
| cmtBReco | -GGAATTCCTTCGGGTAAAAAGGTGC TTC- | *cmtB* promoter region (3077655-636) | PCR fragment for EMSAs | This Study |
| cpxpshort | -GGAATTCCGGGAATGCAGCTCTCGG TC- | *cpxP* promoter region (4103213-3227) | non-specific competitior for EMSAs | Raivio unpublished |

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| cpxuplong | -TTCTGCGGTGACAAGGCGATG CATTG- | *cpxP* promoter region (4103638-3663) | non-specific competitior for EMSAs | Raivio unpublished |
| $D_{61}$E-2 | -GGGCCTGTTGCATAAGTTCTCGCAT CTGCTGA- | *cpxP* sequence (4103598-567) | Used to identify and sequence the λ*plac*Mu53 end | Buelow and Raivio 2005 |
| fecAF | -GACGTTCCCGCTGAAAAATA- | *fecA* promoter region (4514440-421) | PCR fragment for EMSAs | This Study |
| fecAR | -GAGAGGGTAAATCCGCTGTG- | *fecA* promoter region (4514070-089) | PCR fragment for EMSAs | This Study |
| fecIF | -AAAGTGTCCTGGGCAATGTC- | *fecI* promoter region (4515667-686) | PCR fragment for EMSAs | This Study |

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| fecIR | -TGAGAGGCGGTGTAAGTGAA- | *fecI* promoter region (4516240-221) | PCR fragment for EMSAs | This Study |
| Mu3'1 | -AAATTTGCACTACAGGCTTGC- | λp*lac*Mu53 end | Sequencing λp*lac*Mu53 insertions | This Study |
| Mu3'2 | -CGGCATAAGCTGATTTGTGA- | λp*lac*Mu53 end | Sequencing λp*lac*Mu53 insertions | This Study |
| spy-lac3' | -CGGGATCCCGCGGTAGTGGTGTCTGC- | *spy* promoter region (1823564-580) | PCR fragment for EMSAs | Raivio *et al.* 2002 |
| spy-lac5' | -GGAATTCCTTGCTTTCTTATAAATTAATACAG- | *spy* promoter region (1823920-943) | PCR fragment for EMSAs | Raivio *et al.* 2002 |

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| ybiKF | -CTCCTGCGGAGGCAAAAT- | *ybiK* promoter region (865595-612) | PCR fragment for EMSAs | This Study |
| ybiKR | -ACCGCTATGGTCGAGGAC- | *ybiK* promoter region (866277-294) | PCR fragment for EMSAs | This Study |
| ygcLF | -CCAGAGTCATCCCTGCAAAT- | *ygcL* promoter region (2882604-2587) | PCR fragment for EMSAs | This study |
| ygcLR | -GGCGTACAGGGATCCAGTT- | *ygcL* promoter region (2882125-2143) | PCR fragment for EMSAs | This study |
| yicOF | -AACACCGGCAATGTAACGTC- | *yicO* promoter region (3841776-1757) | PCR fragment for EMSAs | This study |
| yicOR | -ACTCACGTAATCGGTATTGTCA- | *yicO* promoter region (3841387-1408) | PCR fragment for EMSAs | This study |

| Primer* | Sequence(5'-3') | Region | Use | Reference |
|---|---|---|---|---|
| yjgXPF | -GGTGTGGTTGGTACGGGATA- | *yjgX* promoter region (4498463-482) | PCR fragment for EMSAs | This Study |
| yjgXPR | -AGCATGCATAACAATTGATCG- | *yjgX* promoter region (4499089-069) | PCR fragment for EMSAs | This Study |
| $W_{30}$C-2 | -AAGTTCTTCACCCGGATGCCAGTTA TCGCCTGAA- | *cpxP* sequence (4103475-508) | Used to identify and sequence the λ*plac*Mu53 end | Buelow and Raivio 2005 |

*Restriction sites incorporated into primers for cloning considerations are indicated by underlining.

#### 2.3.2.1 PCR utilizing random primers

For the purpose of identifying transposon λp*lac*Mu53 insertion points, PCR was performed using the nested primers Mu3'1 and Mu3'2, located at the terminus of the λp*lac*Mu53 insertion, and the random oligonucleotide primers ARB1 and ARB2, which hybridize to repetitive elements in the *E. coli* genome (Caetano-Anolles 1993). The reactions were assembled as described in section 2.3.2, in two stages, using first the ARB1 and Mu3'2 primers. The first reaction was as follows: 1 minute at 95°C followed by six cycles of 30 seconds at 94°C, 30 seconds at 30°C and one minute thirty seconds at 72°C. This was followed by 31 cycles of 30 seconds at 94°C, 30 seconds at 45°C and 2 minutes at 72°C. The reaction was completed with one cycle of 5 minutes at 72°C and 1 minute at 25°C. One μL of this reaction was used as the template for the second reaction mix, which used the primers ARB2 and Mu3'1. The cycles were as follows: 30 cycles of 30 seconds at 94°C, 30 seconds at 55°C and 1 minute 30 seconds at 72°C. These were again followed by one cycle of 5 minutes at 72°C and 1 minute at 25°C. The products of this reaction were analyzed via electrophoresis and DNA sequencing using primer Mu3'1, described in sections 2.3.3 and 2.3.5 respectively.

### 2.3.3 Agarose gel electrophoresis

DNA fragments were separated based on size on 0.8% agarose gels in 1X TAE buffer (40mM Tris (Roche), 1.1mL/L glacial acetic acid, 1mM EDTA (EM Science)). Prior to loading, 1/6 the total volume of loading dye (0.25% bromophenol blue (EM Science), 0.25% xylene cyanol (EM Science), 30% glycerol) was added to each sample.

Twenty μL samples were run alongside a standard size marker (1Kb plus, Invitrogen) for comparison. Gels were run at 100V until the desired level of migration had been attained, the time varied depending on the size of the fragment in question. After electrophoresis, gels were stained in a 1X TAE solution containing ethidium bromide. Bands were visualized using a UV trans-illuminator (Fischer Biotech), and imaged using a FluorChem 5500 imaging system (Alpha Innotech).

### 2.3.4 Purification of DNA from agarose gels

After electrophoresis, the desired products were identified and excised from the gel using a scalpel. The gel slices were transferred to a 1.5mL microfuge tube, and DNA was isolated from the gel according to the manufacturer's instructions for the GeneClean II Kit (Q BioGene).

### 2.3.5 DNA sequencing

DNA sequencing was carried out using DYEnamic™ ET terminator cycle sequencing. The reaction mixture was as follows: 8μL sequencing reaction premix (Molecular Biology Services Unit [MSBU], University of Alberta) 1μL template DNA, either from a plasmid prep or a purified PCR product, 100pM primer, and 7μL $dH_2O$. The sequencing cycle consisted of 25 cycles of 30 seconds at 95°C, 15 seconds at 55°C and 1 minute at 72°C. After the cycling was completed, the reaction products were precipitated by the addition of 2μL sodium acetate (MSBU), and ice cold 95% ethanol, followed by centrifugation at 11,688xg for 15 minutes. The resulting pellet was washed with 200μL

of ice cold 70% ethanol centrifuged at 11,688xg for 2 minutes, then dried overnight. The sample was analyzed by the MSBU, and sequence data was provided in the form of a word document and chromatograph printout.

### 2.3.6 Digestion and cloning of DNA

Restriction digests were carried out according to the instructions provided by the manufacturer of the enzyme (Invitrogen or Roche). The reaction substrate was either 10μL purified PCR product or 10μL of purified plasmid DNA and the reactions were carried out in 35μL volumes. After the reaction was complete, the products were analyzed on an agarose gel to ensure complete digestion, and to allow for purification of the desired product, as per sections 2.3.3 and 2.3.4. Ligation reactions contained an insert to vector ratio of approximately 4:1. Reactions were carried out in a total volume of 10μL, containing 1μL ligase (Invitrogen), 1μL of the suppliers 10x buffer, approximately 6μL DNA and 2μL $dH_2O$ and were incubated overnight at 16°C. The following morning, transformations were carried out as described in section 2.2.2, using 1μL of the ligation reaction, or a 1:10 dilution if the salinity was too high.

## 2.4 A Genetic screen for BaeSR regulated genes

### 2.4.1 λ p*lac*Mu53 infection

An O/N culture of the recipient strain (CA22) was prepared. One millilitre of this O/N culture was mixed with approximately $10^8$ plaque forming units (pfu) of λ p*lac*Mu53

and approximately $10^9$ pfu of λpMu507 helper phage. This mixture was incubated for 30 minutes at 37°C, after which 5mL of LB plus 10mM citrate was added. The culture was then pelleted by centrifugation at RT and 2700xg for 10 minutes and resuspended in 1mL of LB. Serial dilutions were prepared and plated on LB with 100μg/mL kanamycin to select for λp*lac*Mu53 insertion into the bacterial chromosome (Figure 3.1), and 100μg/mL ampicillin to select for pCA14.

### 2.4.2 Isolation and screening of transposition mutants

After the infection mixtures were plated and incubated O/N at 37°C, plates were observed for the formation of colonies (insertion mutants), and plates having a count of single isolate colonies (SICs) in the hundreds were selected for further study. Using a toothpick, individual colonies were patched first onto MacConkey's (Mac) media containing Amp and Kan, and then onto Mac with Amp, Kan, and 0.2% w/v arabinose (ara). These plates were incubated O/N at 37°C and the next morning were scrutinized for any patches which displayed a different *lac* phenotype in the presence of arabinose (and consequently BaeR overexpression). Patches which appeared to be of interest were repatched onto Mac+Amp+Kan, and Mac+Amp+Kan+0.2%Ara, to confirm the phenotype. This streaking was done from the plate containing no arabinose to avoid transferring any arabinose to the non-arabinose plates. Once the phenotype was confirmed, the strain was purified on LB+Amp+Kan, then three SICs were streaked on Mac+Amp+Kan, and Mac+Amp+Kan+0.2%Ara, to confirm the phenotype. Frozen

stocks were prepared of each strain that displayed a consistent phenotype throughout these tests.

### 2.4.3 Confirmation of a BaeR inducible phenotype

The infection protocol using λp*lac*Mu53 phage may produce multiple insertions in the same strain. Therefore it is essential to transduce the insertion into a new strain background to ensure that only one operon fusion is being studied. I utilized P1*vir* transduction to move my insertions into an MC4100 background, selecting for $Kan^R$, as described in section 2.2.3.2. Since these new strains lacked the arabinose inducible BaeR overexpression plasmid pCA14, they were again patched onto Mac+Amp+Kan, and Mac+Amp+Kan+0.2%Ara, this time to ensure that the desired phenotype was lost indicating that it was not due to arabinose induction. Competent cells were prepared from the new strain and transformed with pCA14, as described in sections 2.2.1 and 2.2.2, to test for the reappearance of the arabinose-inducible phenotype. Insertion mutations that did not produce the expected results when subjected to these tests were discarded.

### 2.4.4 Identification of target genes

The locations of the λp*lac*Mu53 insertions were determined by PCR utilizing random primers as described in section 2.3.2.1. The product of the second reaction was examined by gel electrophoresis, and a single band was isolated, purified, and sequenced as described in sections 2.3.3, 2.3.4 and 2.3.5 respectively. The sequence data obtained in this manner were analyzed using the BLAST program available at the Colibri Web

Server, http://genolist.pasteur.fr/Colibri/ (Altschul, 1999). This allowed identification of the gene the insertion was located in or near, and the orientation of the insertion.

## 2.5 MBP-BaeR protein purification and analysis

### 2.5.1 Preparation of whole cell lysates

A 5ml O/N culture of a bacterial strain containing pMAL-c2:*baeR* (AC31) was prepared. The next day, 5mL of this culture was added directly into 1L of LB+ 0.2% glucose, and Amp. This was grown to an $OD_{600}$ of approximately 0.5. At this point a 1mL sample was taken, harvested by centrifugation at 11,680xg for one minute and resuspended in 50μL of 2X SDS-PAGE buffer (125mM Tris-Cl, 20% glycerol, 10% β-mercaptoethanol, 6% SDS, and 2% Bromophenol Blue). Three millilitres of 0.1M isopropyl-beta-D-thiogalactopyranoside (IPTG) (Sigma) were added to the growing culture in order to induce MBP-BaeR expression from the plasmid pMAL-c2:baeR, and the culture was permitted to grow for an additional 2 hours. A second sample of 0.5mL was taken and the sample was pelleted by centrifugation at 11,680xg for one minute and resuspended in 100μL of 2X SDS-PAGE sample buffer. Twenty microlitres of each of these samples were later analyzed by electrophoresis on a 12% polyacrylamide gel stained with Coomassie Blue (as described in section 2.5.5) to ensure that MBP-BaeR induction had proceeded as expected. The culture was divided into 4x250mL aliquots which were centrifuged at 4000xg for 20 minutes. The supernatant was then decanted from the pellets, and these were placed at -20°C overnight. The pellets were subsequently

thawed on ice, and resuspended in 50mL column buffer (0.02M Tris-Cl pH 7.4, 0.2M NaCl (EMD), 0.001M EDTA, and 0.001M phenylmethylsulphonalfluoride (Sigma)). The cells were lysed by three sequential passes through a French Press (American Instrument Company). A soluble fraction was prepared by sedimenting the lysate at 9000xg for 30 minutes. The supernatant was reserved, and stored at -20°C.

### 2.5.2 MBP-BaeR protein purification using affinity chromatography

Affinity columns were prepared by placing 20mL of Amylose resin (NewEngland Biolabs) into a 20mL syringe (BD) on a bed of compacted glass wool. The column was washed with 8 volumes of column buffer, at which point the soluble extracts were loaded onto the column. After this, the column with bound MBP-BaeR was washed with 12 volumes of column buffer to remove any contaminating cellular proteins. The bound MBP-BaeR was eluted using column buffer with 10mM maltose (Sigma), and 25x 3mL samples were collected using a fraction collector (Cygnet). Protein assays were performed as described in section 2.5.3 to determine which fractions contained the eluted MBP-BaeR protein. These fractions were pooled, and purity was assayed by combining 10μL of the pooled protein fraction with 10μL of 2x SDS-PAGE sample buffer and performing SDS-PAGE as in section 2.5.5. If the affinity column was to be reused at a later point, it was preserved in column buffer with 0.02% sodium azide. The resin was regenerated by the following washes: 3 column volumes of $dH_2O$, 3 column volumes of 0.1% SDS, 1 column volume of $dH_2O$, and 5 column volumes of column buffer.

### 2.5.3 Bradford assay for protein concentration

Protein concentration was determined using the Bradford assay according to the manufacturer's instructions (Biorad). The BSA standards used were 1.5, 1.0, 0.8, 0.5 and 0.2mg/mL.

### 2.5.4 Protein buffer exchange

The buffer that the MBP-BaeR protein was in was changed from column buffer to PBS (140mM NaCl, 10mM $Na_2HPO_4$ (BDH), 3mM KCl (BDH), and 2mM $KH_2PO_4$ (BDH), pH 7.4) using Econo-Pac® 10DG Columns (Biorad), according to the manufacturer's instructions.

### 2.5.5 SDS-polyacrylamide gel electrophoresis

Protein samples were analyzed via SDS-polyacrylamide gel (PAGE) using a two gel system. The separating gel contained 12% acrylamide, and the stacking gel consisted of 5% acrylamide (Laemmli 1970). Gels were run for one hour at 100V, in Tris-Glycine running buffer (25mM Tris, 157mM glycine (Fischer Scientific), and 0.1% SDS).

## 2.6 Electrophoretic mobility shift assays

### 2.6.1 Preparation of end-labeled DNA fragments

DNA fragments containing the promoter regions of target genes were amplified and purified as described in sections 2.3.1 and 2.3.3 respectively. The purified products were quantified by measuring the absorbance at 260 nm using a Ultraspec 3000 spectrophotometer (Pharmacia Biotech). Five picomoles(pmol) of each fragment were then radioactively end-labeled with $\gamma^{32}$P-dATP using T4 polynucleotide kinase (Invitrogen) according to the manufacturer's instructions. In order to remove any unincorporated label, size exclusion chromatography was performed using Probe Quant™ G50 Micro Columns (Amersham Pharmacia), according to the manufacturer's instructions, resulting in a final sample volume of 50μL.

### 2.6.2 Phosphorylation of MBP-BaeR by a small phospho-donor

Phosphorylation reactions were prepared in a 9μL volume containing 1X binding buffer (10mM Tris-Cl, 50mM KCl, 25mM $MgCl_2$, 1mM dithiothreitol, 2.5% glycerol, 0.05%NP-40, and 50ng poly[dI-dC]), 50mM acetyl phosphate (lithium potassium salt, Sigma), and varying concentrations, one each of 0μM, 0.8μM, 1.0μM, 1.5μM, and 2.0μM, of MBP-BaeR. These reactions were incubated at 37°C for 30 minutes to allow phosphorylation to occur.

### 2.6.3 Electrophoretic mobility shift assay

One microliter of purified labeled promoter fragment (section 2.6.1) was added to the phosphorylated MBP-BaeR (section 2.6.2). Control reactions were prepared by adding either 1μL of purified unlabeled *cpxP* promoter region or 1μL of a purified

unlabeled sample of the promoter region under study to redundant tubes containing the radio-labeled probe and the highest concentration of protein prepared as described above. Either control would represent an excess of cold competitor, with the *cpxP* promoter serving as a control for non-specific binding, and the unlabeled promoter under study serving as a control for specific binding. Binding was allowed to occur for 30 minutes at room temperature, after which 4μL of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) was added. Samples were run on a non-denaturing polyacrylamide slab gel (10% acrylamide, 0.1% ammonium persulfate (APS), 1X Tris-acetate buffer) in 1xTAE buffer at 100V for approximately 1.5 hours.

### 2.6.4 Autoradiography

After electrophoresis, the gel was placed onto a labeled piece of #3 Whatman® paper and dehydrated using a model 583 gel dryer (Biorad, courtesy the Leskiw lab, University of Alberta) set to 80°C for 45 minutes. The dried gel was encased in plastic wrap, and placed into a storage phosphor screen (Molecular Dynamics, courtesy the Frost lab, University of Alberta). The screen was exposed overnight, then developed the following day.

## 2.7 DNaseI protection assay

### 2.7.1 Preparation of end-labeled DNA fragments

In order to produce PCR products radio-labeled on the 5' end of a single strand, radio-labeled primers were produced prior to the PCR reaction. 100pmol of the primer for the strand to be labeled was labeled with $\gamma^{32}$P-dATP using T4 polynucleotide kinase (Invitrogen) according to the manufacturer's instructions. The primer was then precipitated by adding 1μL glycogen (Roche) and 110μL of ice cold ethanol followed by an overnight incubation. The primers were then pelleted by centrifugation at 4024xg for 10 minutes, dried at RT until no trace of ethanol was observed, and resuspended in 16μL of $H_2O$. PCR was then carried out as detailed in section 2.3.2.

### 2.7.2 Purification of radio-labeled DNA from polyacrylimide gels

In order to purify the previously created radio-labeled PCR products, polyacrylamide gel electrophoresis was carried out as described as described in section 2.6.3. The resulting gel was transferred to a solution of 1xTAE buffer containing Ethidium bromide, allowed to stain for 20 minutes. Following this, the DNA was visualized under UV illumination, and the desired bands were excised. DNA was eluted from the gel by being crushed using a pipette tip in 400μL elution buffer (0.5M ammonium acetate, 1mM EDTA p.H. 8.0), and incubation at 37°C O/N. The buffer was removed to a new centrifuge tube, and the gel was washed with 200μL elution buffer which was then added to the first aliquot of buffer. The eluted DNA was then precipitated

by the addition of 2μL glycogen and 900μL of ice cold ethanol. The DNA was pelleted by centrifugation at 4024xg for 10 minutes at room temperature. The pellet was washed with 500μL of ice cold ethanol, and then allowed to dry prior to resuspension in 20μL $H_2O$.

### 2.7.3 DNaseI Protection assay

Five centrifuge tubes were assembled, containing 3μL radio-labeled DNA (section 2.7.1), 5mM MgCl, 10% glycerol, 1x binding buffer (as in section 2.6.2) and 0, 0.6, 0.9, 1.2, or 1.5μM MBP-BaeR in a final volume of 20μL. Binding was allowed to occur for 30 minutes. To these were added 1μL of DNaseI (0.025U/μL), after which, digestion was allowed to proceed for 30sec followed by the addition of 50μL buffer saturated phenol to stop the reaction. Thirty microlitres of $H_2O$ was added to bring the aqueous layer to 50μL, and the microfuge tube was vortexed for 10 seconds. The tube was the centrifuged for 2 minutes at 4024xg and the aqueous layer was removed to a new microfuge tube. The extraction procedure was then repeated using phenol:chloroform followed by chloroform. After the final extraction, the DNA was precipitated by the addition of 1μL glycogen, 200μL ice cold ethanol and incubation on ice for 30 minutes. The DNA was pelleted by centrifugation for 10 minutes at 4024xg and dried at RT to remove any remaining ethanol. The pellet was resuspended in 4μL formamide loading dye (1mL formamide, 10μL 0.5M EDTA, 0.0025g bromophenol blue, and 0.0025g xylene cyanol).

### 2.7.4 Sanger sequencing protocol

To identify the sequence of the DNA protected by BaeR binding in the DNaseI protection assays, DNA sequencing was carried out using a Thermo Sequenase radiolabeled terminator cycle sequencing kit (USB) according to the manufacturer's instructions. For an individual DNaseI protection assay, sequencing would be carried out using the same primer used to generate the probe as described in 2.7.1. The template was a purified non-labled PCR fragment of the promoter under study, created using primers AcrD-F-ER1 and AcrD-R-BH1.

### 2.7.5 Polyacrylamide gel electrophoresis for sequencing

In order to visualize the results of the DNaseI protection assays, polyacrlamide gel electrophoresis was performed using a Sequi-Gen GT Nucleic Acid Electrophoresis cell (Bio-Rad). Sequencing gels were run at a constant 40W for 2 hours in 1xTBE buffer (10x Stock: 54g tris base, 27.5g boric acid, 20mL 0.5M EDTA p.H. 8.0). The gel was as follows, 50mL 6% sequencing solution (230g urea, 50mL 10xTBE buffer and 100mL 30% acrylimide brought to 500mL in $H_2O$), and 187.5μL 10% APS. Prior to loading, the gel was prerun at 40W until it had reached a temperature of 45°C. When this temperature was achieved, the samples were heated to 80°C then cooled on ice for 2 minutes prior to loading alongside the sequencing ladder. The gel was allowed to run for 1 hour and 30 minutes, then dried for 2 hours and subjected to autoradiography as detailed in section 2.6.4.

## 2.8 Reporter Assays

### 2.8.1 β-Galactosidase assay

β-Galactosidase assays were performed to examine the expression of the *lacZ* operon fusions in a variety of genetic backgrounds following the protocol of Slauch and Silhavy (1991). Briefly, O/N cultures were prepared in quintuplicate for each strain to be assayed. 1:50 subcultures were prepared and allowed to grow for three hours, after which induction was carried out for an additional 2 hours if chemical inducing cues were being utilized. The cells were harvested by centrifugation at RT and 2700xg for 10 minutes and resuspended in 2mL of 1x Z-buffer (60mM $Na_2HPO_4 \cdot 7H_2O$, 40mM $NaH_2PO_4 \cdot H_2O$, 10mM KCl, 1mM $MgSO_4 \cdot 7H_2O$, 2.7μL β-mercaptoethanol/mL). The $OD_{600}$ was taken fron 250μL samples of each tube placed into a microtiter dish. Two drops of chloroform and one drop of 1% SDS was added to the remaining cells and lysis was achieved by vortexing for 30 seconds. 50μL of each sample was combined with 150μL 1x Z-buffer and 50μL fresh 10mg/mL ONPG in a 96 well microtiter dish and the $\Delta A_{420}$ was monitered over a period of 30 minutes. The $OD_{600}$ and $A_{420}$ were determined using a Wallac Victor2 1420 multilabel counter. The resulting data was converted into β-Galactosidase units via the following formula $MU=(\Delta A_{420}/\text{time} * 1000 * 60)/ A_{600}$ . For each strain, outlying values were removed, and the remaining three data points were averaged and plotted with their standard deviation. Comparisons were made between strains in non-inducing conditions and inducing conditions. The inducing cues used in this study were a genetic background containing the gain-of-function allele

*baeS1*::Tn*10cam*, which overexpresses the BaeSR pathway, the presence of the established chemical inducer 2mM indole, and the novel chemical inducers, 10 μg/mL myricetin, and 10mg/mL $NaWO_4$.

# 3. Characterization of Novel BaeSR regulated genes

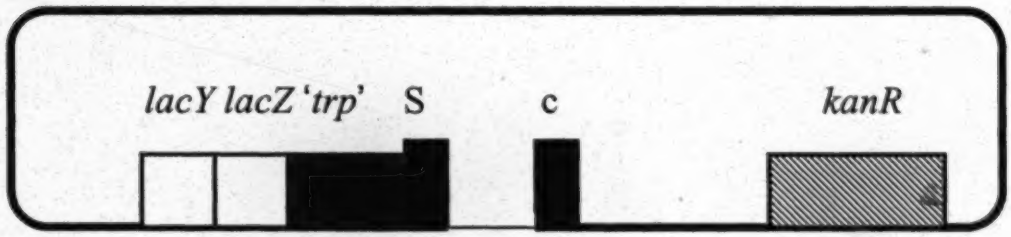
## 3.1 Genes identified by transposon screen

At the start of this project, only three promoters were known to be controlled by the BaeSR regulon, those of *mdtABCDbaeSR, acrD*, and *spy*, all of which are related to stress response. However, it had been observed that the established BaeR activating signal indole was utilized as a quorum sensing molecule, activating the expression of a set of genes involved in the catabolism of amino acids as it accumulates during stationary phase (Wang *et al.* 2001). This raised the possibility that the BaeSR system may be involved in the indole dependent regulation of genes beyond those associated with the envelope stress response (Raffa, and Raivio 2002). This project was intended to further clarify the physiological role of the BaeSR pathway by expanding our knowledge of the regulon it controls, identifying additional genes which may or may not play a role in envelope stress response. To this end, we undertook a transposon screen using the λp*lac*Mu53 transposon to screen for altered gene expression in the presence of BaeR overexpression. λp*lac*Mu53 is a randomly integrating transposon which integrates into the bacterial chromosome using the termini of the Mu phage (Figure 3.1). It contains a promoterless *lacZ* reporter gene adjacent to the initiation codon of the *trpA* gene and when inserted into a chromosomal operon in the correct orientation produces a TrpA-LacZ hybrid protein under the control of the native chromosomal promoter (Bremer *et al.*

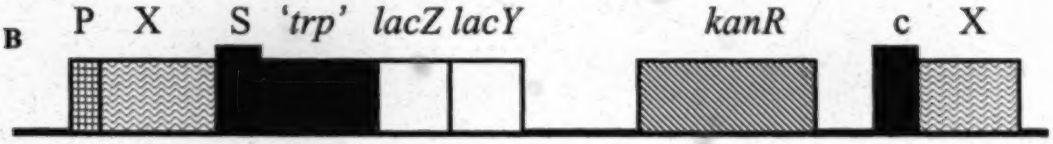
**Figure 3.1. Schematic diagram of λp*lac*Mu53 insertion in the bacterial chromosome.** **A) λp*lac*Mu53**. The thick lines represent λ DNA, while the thin line is bacterial DNA, originating from the previous host strain insertion point. The bacteriophage Mu termini are represented by black boxes labled S and c, the striped box is a kanamycin resistance cassette, the gray box is a portion of the *trp* operon, and the white boxes are promoterless *lacZY* genes.

**B) Operon Fusion.** λp*lac*Mu53 as an operon fusion in the bacterial chromosome. Legend as above, with X representing the interrupted operon, and P representing its promoter. Adapted from Weinstock *et al.* 1985.

A




B



1985). Strains carrying an insertion mutation may be selected for using a medium containing kanamycin. We screened for operon fusions whose expression was altered by BaeR overexpression as this would be expected to lead to constitutive expression of regulon members (or repression of regulon members downregulated by BaeR activation) since overexpression of response regulators has been shown to mimic constitutive pathway activation. This is possibly because a small portion of the response regulator pool is known to be phosphorylated in the absence of activation of its cognate histidine kinase (Aguilar *et al.* 2001, Powell and Kado, 1990). BaeR overexpression was achieved by the use of plasmid pCA14, which is a pBAD18 derivative, carrying *baeR* downstream of the promoter $Para_{BAD}$, which is strongly inducible by 0.2% arabinose.

In order to generate insertion mutation strains for study, λp*lac*Mu53 infection was carried out on *E.* coli strain CA22, an *E. coli* MC4100 derivative carrying BaeR overexpression plasmid pCA14, and the infection mixture was plated on ampicillin and kanamycin to select for the plasmid and insertion mutations. Single isolated colonies (SICs) resulting from this procedure were screened for phenotypic differences on Mac and Mac + 0.2% ara. We screened a total of 3728 transposition mutants for altered *lacZ* expression in the presence of BaeR overexpression. Patches which showed a phenotype of interest, either increased or decreased β-Galactosidase activity in the presence of arabinose and BaeR overexpression, were repurified on LB+kan+amp, and SICs were again tested on Mac and Mac+ara to ensure a consistent phenotype.

Five strains which showed an alteration in their *lac* phenotype in the presence of arabinose and BaeR overexpression were selected for further study, three of which

appeared to be up-regulated by BaeR, and two of which appeared to be downregulated (data not shown). These insertion mutations were transduced by P1*vir* transduction into *E. coli* MC4100 to create stains CO4, CO2, CO8, CO23, and CO16. This was done in order to ensure that only a single copy of λp*lac*Mu53 was present, as multiple insertions are possible, and to allow us to verify that an arabinose inducible phenotype was not observed by looking for an absence of the 0.2% ara inducible phenotype in the absence of ara inducible BaeR overexpression plasmid pCA14. None of these strains exhibited an arabinose inducible phenotype (data not shown). Once this was confirmed, these strains were transformed with pCA14 to create strains CO67, CO47, CO10, CO24, and CO46 and the return of the BaeR dependent phenotype was established (Figure 3.2, 3.3). We observed that when these strains were streaked on Mac media in the presence and absence of 0.2% ara, strains CO67, CO47, and CO10 had increased β-galactosidase production on Mac+0.2% ara, while strains CO24 and CO46 had decreased β-galactosidase production on Mac+0.2% ara (Figure 3.2, 3.3). This indicates that these strains are indeed responding to BaeR over production, rather than arabinose.

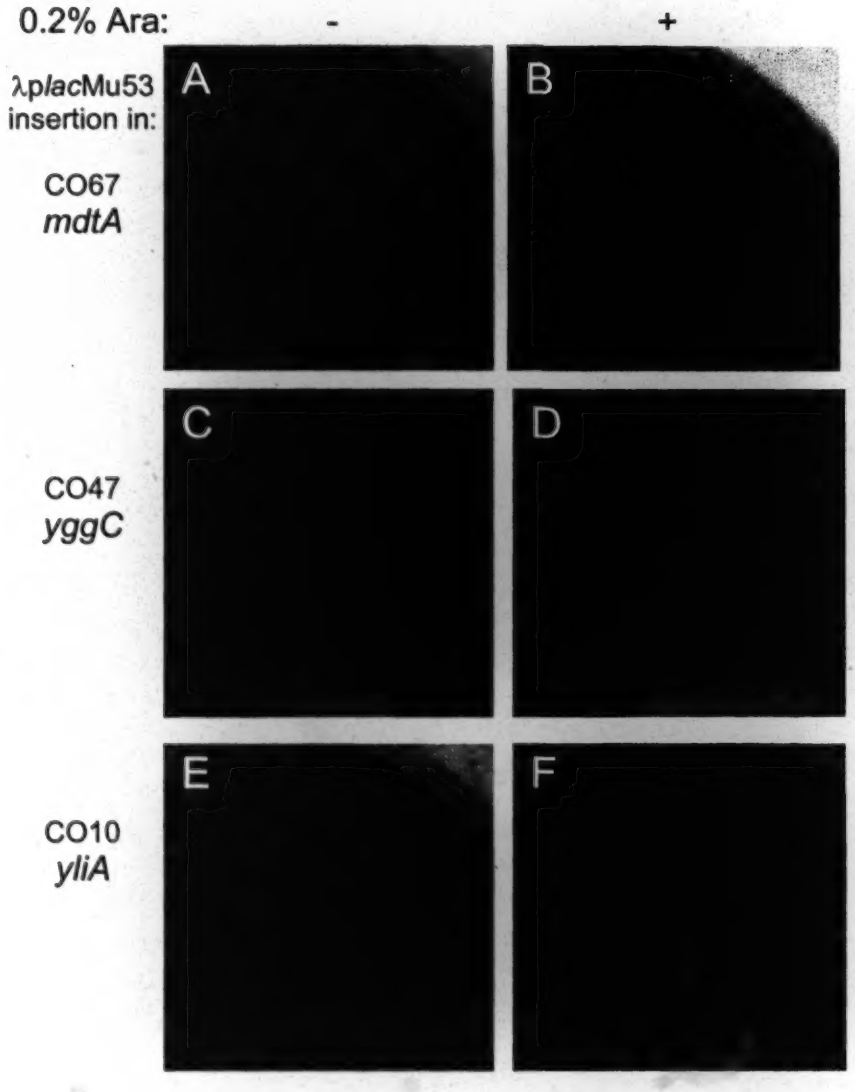
At this point the identities of the genes containing λp*lac*Mu53 insertions were determined using random oligonucleotide PCR, with one primer located in the Mu phage terminus, and the other binding randomly to repetitive genome elements as described in section 2.3.2.1 (Figure 3.4). Many of the insertions were found to be located in genes that appeared to be members of operons based on their context on the chromosome. Strain CO4 was determined to contain a λp*lac*Mu53 insertion in *mdtA* at position 2153145bp of the *E. coli* genome. This corresponds to a location 1107 bp downstream from the *mdtA*

**Figure 3.2. λp*lac*Mu53 insertions in *mdtA, yggC* and *yliA* are upregulated by BaeR overexpression.**

P1*vir* lysates were prepared from *E. coli* strains hosting λp*lac*Mu53 insertions identified in our screen as being upregulated in the presence of BaeR overexpression and used to transduce the insertions into *E. coli* MC4100. The resulting strains were then transformed with pCA14, which overexpresses BaeR in the presence of 0.2% arabinose. These strains were named CO67, CO47 and CO10 and contain λp*lac*Mu53 insertions in *mdtA, yggC* and *yliA* respectively. We observe that in all strains, expression of the *lacZ* reporter is increased when SICs of each strain are streaked out on Mac + 0.2% ara, resulting in BaeR overexpression as compared to when they are streaked out on Mac.

A) CO67 on Mac, B) CO67 on Mac + 0.2% ara, C) CO47 on Mac,

D) CO47 on Mac + 0.2% ara, E) CO10 on Mac, F) CO10 on Mac + 0.2% ara.




MacConkey's Medium

**Figure 3.3. λp*lac*Mu53 insertions in *fecD* and *yjgX* are downregulated by BaeR overexpression.**

P1*vir* lysates were prepared from *E. coli* strains hosting λp*lac*Mu53 insertions identified in our screen as being downregulated in the presence of BaeR overexpression and used to transduce the insertions into *E. coli* MC4100. The resulting strains were then transformed with pCA14, which overexpresses BaeR in the presence of 0.2% arabinose. These strains were named CO24, and CO46 and contain λp*lac*Mu53 insertions in *fecD,* and near *yjgX* respectively. We observe that in each strain, expression of the *lacZ* reporter is decreased when SICs of each strain are streaked out on Mac + 0.2% ara, resulting in BaeR overexpression as compared to when they are streaked out on Mac.

A) CO24 on Mac, B) CO24 on Mac + 0.2% ara, C) CO46 on Mac,

D) CO46 on Mac + ara.

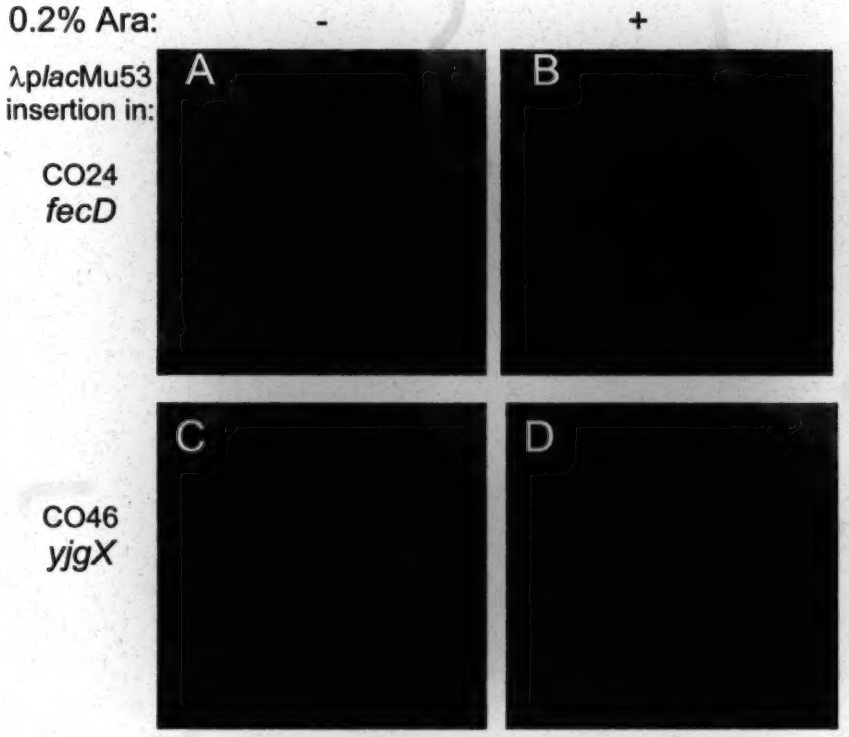
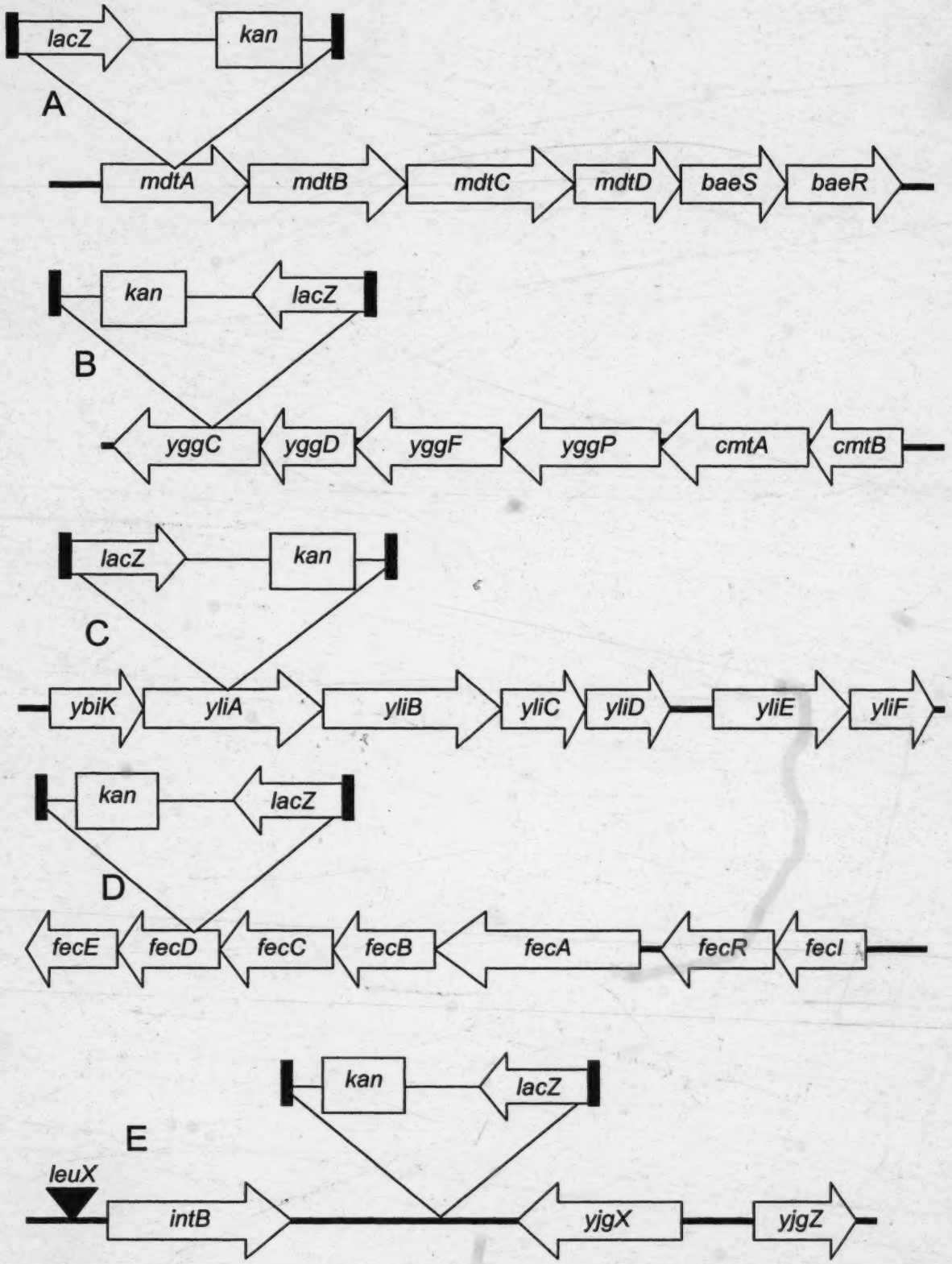

0.2% Ara:
-
+
λplacMu53 insertion in:
CO24
fecD
CO46
yjgX
A
B
C
D
MacConkey's Medium

**Figure 3.4. Chromosomal insertion points of the λp*lac*Mu53 mutations**

The insertion points of the λp*lac*Mu53 insertions were determined by a two step PCR process utilizing the random oligonucleotides ARB1 and ARB2. Colony PCR was performed first using primers ARB1 and Mu3'2. A 1μL sample of this reaction was then used as template for a second PCR reaction using nested primers ARB2 and Mu3'1. This reaction was run out on an agarose gel, bands were excised and purified, then sequenced using primer Mu3'1.

These diagrams depict the locations and directions of the λp*lac*Mu53 insertions isolated in this study. The precise insertion locations relative to the *E. coli* genome are as follows:

A) λp*lac*Mu53 insertion in *mdtA* at position 2153145bp, 1107bp from the *mdtA* start site. B) λp*lac*Mu53 insertion in *yggC* at position 3072208bp, 501bp from the *yggC* start codon. C) λp*lac*Mu53 insertion in *yliA* at position 867570bp, 845bp away from the *yliA* start codon, D) λp*lac*Mu53 insertion in *fecD* at position 4509814, 168bp away from the *fecD* start site. E) λp*lac*Mu53 insertion near *yjgX* at position 4495821bp, 2538bp away from the *yjgX* start codon.


lacZ
kan
A
mdtA
mdtB
mdtC
mdtD
baeS
baeR
kan
lacZ
B
yggC
yggD
yggF
yggP
cmtA
cmtB
lacZ
kan
C
ybiK
yliA
yliB
yliC
yliD
yliE
yliF
kan
lacZ
D
fecE
fecD
fecC
fecB
fecA
fecR
fecI
kan
lacZ
E
leuX
intB
yjgX
yjgZ

start site. The identification of this gene by our screen confirms that our screening procedure was able to identify insertions in genes which are known to be regulated by the BaeSR pathway. The λp*lac*Mu53 insertion in CO2 was found to be located in the *yggC* gene in position 3072208bp of the genome, corresponding to a location 501bp downstream from the *yggC* start codon. *yggC* is thought to be a member of an operon consisting of *cmtBAyggPFDC*. Though its function has not been experimentally determined, *yggC* is thought to be a kinase based on sequence homology. Of the other members of the operon, again based on sequence homology, *yggD* appears to be a transcriptional regulator, *yggF* has homology to the fructose 1,6-bisphosphatase *glpX*, *yggP* has homology to the *Klebsiella pneumonia* oxidoreductase *sorE* and *cmtBA* appears to encode a phosphoenolpyruvate-dependent mannitol phosphotransferase system. Strain CO8 was found to harbour a λp*lac*Mu53 insertion in *yliA*, located 845 bp downstream of its start codon, in position 867570bp of the genome. *yliA* is believed to be a part of an operon consisting of *ybiKyliABCD*. The *yliA* gene is believed be the ATPase responsible for energy coupling in an ABC type transporter, the remaining members of which are encoded by *yliBCD*. *ybiK* is thought to be an L-asparaginase precursor of the asparaginase 2 family, which would undergo autoproteolysis upon maturation (Borek *et al.* 2004). The λp*lac*Mu53 insertion in CO23 was found to be located at 4509814bp of the genome, in *fecD*, a member of the *fecABCDE* operon. The insertion point is 168bp downstream of the *fecD* start codon. This operon constitutes the iron (III) dicitrate transport system of *E. coli*, with *fecCD* representing the inner membrane permease, *fecE* encoding an ATP binding protein, and *fecB* and *fecA* encoding the periplasmic and outer

membrane components respectively. Lastly, the λp*lac*Mu53 insertion in strain CO16 was found to be located at 4495821bp in the genome, which places it in an intergenic region adjacent to and in the same orientation as the *yjgX* gene, 2538bp downstream of its start codon. *yjgX* has homology to sulfatases and may be an integral membrane protein.

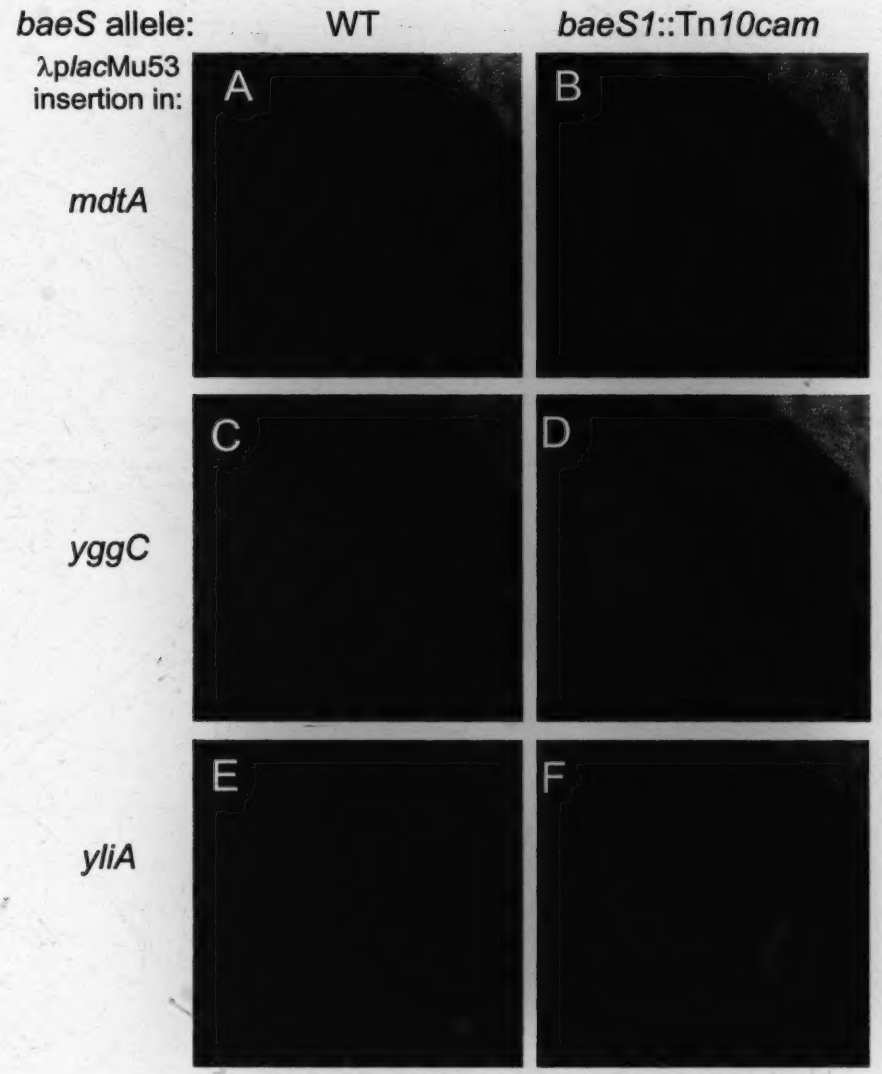
### 3.2 Induction of operon fusions by the *baeSl*::Tn*10cam* allele

In order to support our hypothesis that our screen had identified genes which are members of the BaeSR pathway, P1*vir* transduction was used to move the λp*lac*Mu53 insertions in *mdtA, yggC, yliA, fecD*, and near *yjgX* into strain TR776, which contains the *baeS1::*Tn*10cam* allele in an *E. coli* MC4100 background, creating strains CO63, CO58, CO59, CO75, and CO61 respectively. The *baeS1::*Tn*10cam* allele is a gain of function mutation resulting from the insertion of a chloramphenicol resistant mini-Tn10 derivative into the fifth codon of *baeS,* resulting in increased activation of the BaeSR pathway (Raffa and Raivio 2002). We hypothesized that the presence of this allele would confer a phenotype similar to that of BaeR overexpression on strains bearing the λp*lac*Mu53 insertions. These strains were plated on Mac media next to strains bearing the λp*lac*Mu53 insertions in a wild type MC4100 background in order to assay differences in β-galactosidase production. We observed that for strains bearing the λp*lac*Mu53 insertions in *mdtA, yggC* and y*liA,* which were upregulated by BaeR overexpression (Figure 3.2), the *baeS1*::Tn*10cam* allele caused a noticeable increase in β-galactosidase production (Figure 3.5). For strains bearing the λp*lac*Mu53 insertions in *fecD* and near

**Figure 3.5. λp*lac*Mu53 insertions in *mdtA, yggC and yliA* are activated by the *baeS1::*Tn*10cam* allele.**

P1*vir* lysates were prepared from strains hosting λp*lac*Mu53 insertions identified in our screen as being upregulated in the presence of BaeR overexpression and used to transduce the insertions into *E. coli* strains MC4100 (wild type) and TR776 (MC4100 *baeS1::*Tn*10cam*). We observed that expression of *lacZ* reporters in *mdtA* (strains CO4, CO63), *yggC* (strains CO2, CO58), and *yliA* (strains CO8, CO59) was increased in the presence of the *baeS1* gain of function mutation (Compare A, C, and E with B, D, and F respectively) when single isolated colonies from each of these strains were streaked out on Mac.

A) CO4, λp*lac*Mu53 in *mdtA*, MC4100 background B) CO63, λp*lac*Mu53 in *mdtA*, TR776 background C) CO2, λp*lac*Mu53 in *yggC*, MC4100 background D) CO58, λp*lac*Mu53 in *yggC*, TR776 background E) CO8 λp*lac*Mu53 in *yliA*, MC4100 background F) CO59 λp*lac*Mu53 in *yliA* TR776 background.


baeS allele:
WT
baeS1::Tn10cam
λplacMu53 insertion in:
mdtA
yggC
yliA
A
B
C
D
E
F

*yjgX*, which were downregulated (Figure 3.3), the presence of the *baeS1::*Tn*10cam* allele resulted in a decrease in β-galactosidase production, though this was not visually evident for the λp*lac*Mu53 insertion in *fecD* (Figure 3.6). These observations provide support for the hypothesis that the genes these λp*lac*Mu53 insertions are located in do indeed respond to the BaeSR pathway. In order to provide a quantitative basis for comparison using these strains, we performed β-galactosidase assays on them in the mid-log, and stationary phases of growth. O/N cultures were prepared for each strain in quintuplicate, and subsequently subcultured 1:50 and allowed to grow for three hours. At this point a β-galactosidase assay was performed using the procedure of Slauch and Silhavy (1991). The $OD_{600}$ and $A_{420}$ were determined and β-galactosidase activity was calculated in β-galactosidase units using the formula $MU=(\Delta A_{420}/time * 1000 * 60)/ OD_{600}$. For each strain, up to two outlying values were discarded, and the remaining data points were averaged and the standard deviation was determined. Figure 3.7 displays the results of this assay, and confirms the results of the plates shown in Figures 3.5 and 3.6. The results observed were similar between the mid-log and stationary phase cultures, with a tendency for the expression levels in the stationary phase samples to be higher, possibly reflecting growth phase dependent regulation. In mid-log phase cultures expression of β-galactosidase from strains carrying the λp*lac*Mu53 insertions in *mdtA, yggC,* and *yliA*, whose expression was expected to be increased by the *baeS1::*Tn*10cam* allele, was found to be increased 2.3, 7.7, and 60.8 fold respectively. It is striking that the newly identified genes *yggC* and *yliA* give stronger induction than the previously characterized *mdtA,* however this result should be considered in light of the very low levels of expression seen

**Figure 3.6. λp*lac*Mu53 insertions in *fecD* and *yjgX* are downregulated by the *baeS1*::Tn*10cam* allele.**

P1*vir* lysates were prepared from strains hosting λp*lac*Mu53 insertions identified in our screen as being downregulated in the presence of BaeR overexpression and used to transduce the insertions into *E. coli* strains MC4100 (wild type) and TR776 (MC4100 *baeS1*::Tn*10cam*). Expression of *lacZ* reporters in *fecD* (strains CO23, CO75), and near *yjgX* (strains CO16, CO61) decreased in the presence of the *baeS1* gain of function mutation (Compare A, and C with B, and D) when SICs from each of these strains are streaked out on Mac.

A) CO23, λp*lac*Mu53 in *fecD*, MC4100 background B) CO75, λp*lac*Mu53 in *fecD*, TR776 background C) CO16, λp*lac*Mu53 near *yjgX*, MC4100 background D) CO61, λp*lac*Mu53 near *yjgX*, TR776 background.

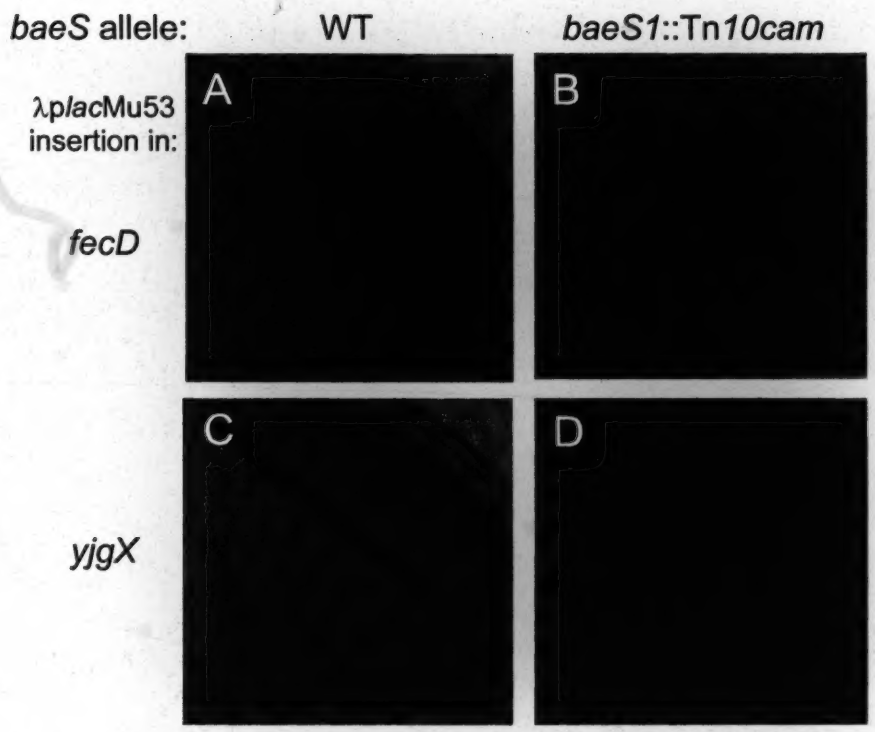
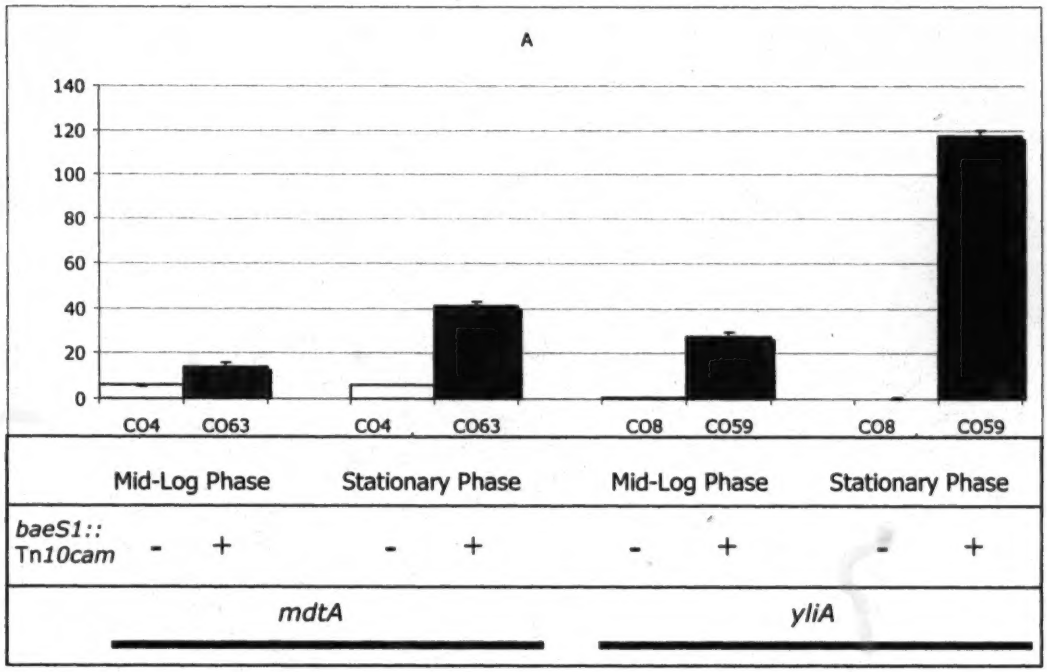

baeS allele:
WT
baeS1::Tn10cam
λplacMu53 insertion in:
fecD
yjgX
A
B
C
D

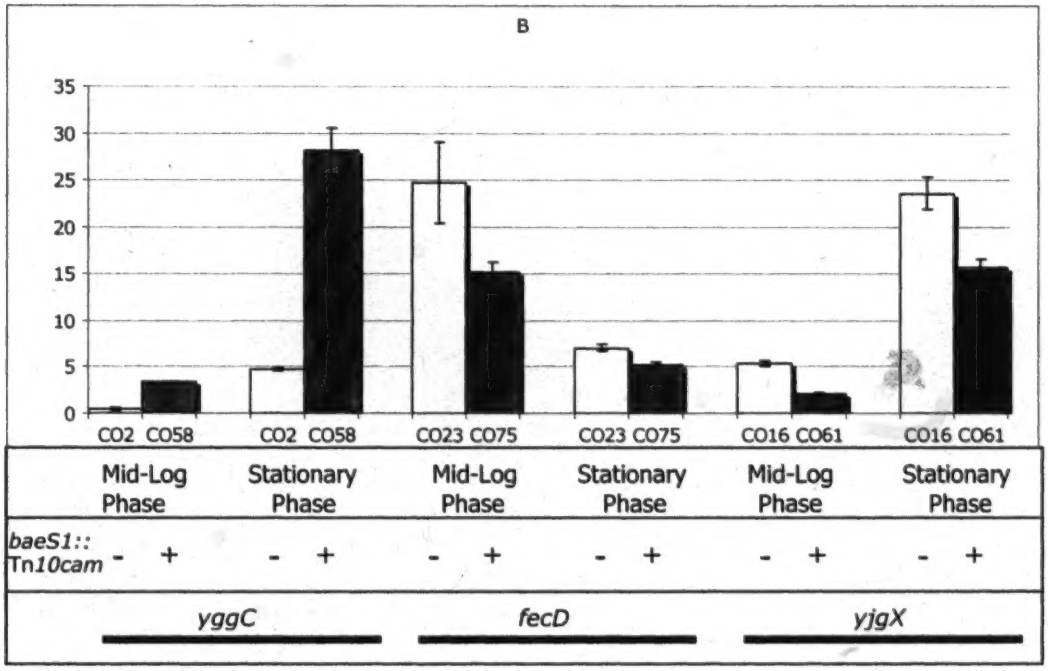
**Figure 3.7. β-galactosidase assay showing the affect of the *baeS1*::Tn*10cam* allele upon expression of β–galactosidase by the λp*lac*Mu53 insertion mutations.**

Five samples of each strain were grown overnight (stationary phase), then subcultured 1:50 and allowed to grow for three hours at 37°C (mid-log phase). At this point a β–galactosidase assay was performed as described in Slauch and Silhavy (1991). The $OD_{600}$ and $A_{420}$ were determined and converted into β-Galactosidase units using the formula $MU=(\Delta A_{420}/time * 1000 * 60)/ A_{600}$. These units appear on the Y axis of the graphs. For each strain, up to two outlying values were discarded, and the remaining data points were averaged and the standard deviation was determined. This experiment was replicated with similar results (not shown).

A) β–galactosidase production by strains carrying λp*lac*Mu53 insertions in *mdtA* (left), and *yliA* (right), in the presence (gray bar) or absence (white bar) of the *baeS1*::Tn*10cam* allele.

B) β–galactosidase production by strains carrying λp*lac*Mu53 insertions in *yggC* (left), *fecD* (middle), and near *yjgX* (right), in the presence (gray bar) or absence (white bar) of the *baeS1*::Tn*10cam* allele.


A
140
120
100
80
60
40
20
0
CO4 CO63 CO4 CO63 CO8 CO59 CO8 CO59
Mid-Log Phase
Stationary Phase
Mid-Log Phase
Stationary Phase
baeS1::
Tn10cam - + - + - + - +
mdtA
yliA


B
35
30
25
20
15
10
5
0
CO2 CO58 CO2 CO58 CO23 CO75 CO23 CO75 CO16 CO61 CO16 CO61
Mid-Log Phase
Stationary Phase
Mid-Log Phase
Stationary Phase
Mid-Log Phase
Stationary Phase
baeS1::
Tn10cam - + - + - + - + - + - +
yggC
fecD
yjgX

for these genes in the absence of the *baeS1::*Tn*10cam* allele, compared to the higher level of basal expression observed for *mdtA*. We expected mid-log phase cultures of strains bearing the λp*lac*Mu53 insertions located in *fecD* and near *yjgX*, to exhibit downregulated β-galactosidase production in the presence of the *baeS1::*Tn*10cam* allele. The fold changes in β-galactosidase production in these strains in response to the presence of the *baeS1::*Tn*10cam* allele are more modest, at 0.6 and 0.4 fold respectively. Again, these results corroborate those obtained in our BaeR overexpression studies.

### 3.3 Induction of operon fusions via inducers of the BaeSR pathway

We performed β-galactosidase assays in order to determine if we could induce expression of our λp*lac*Mu53 insertions using known inducers of the BaeSR pathway. The inducers chosen for this study were indole, which is believed to cause membrane derangement, and has been shown to induce expression of the *spy* gene in a BaeSR dependent manner (Garbe *et al.* 2000, Raffa and Raivio 2002). We also looked for induction by two novel inducers of the pathway, myricetin, and $NaWO_4$. Myricetin, which is a plant flavonoid shown to have activity against antibiotic resistant bacteria, and $NaWO_4$, a toxic analog of molybdate, were found to repress the growth of a BaeSR null strain to a greater extent than a wildtype strain (Xu and Lee 2001, Rech *et al.* 1996, Zhou *et al.* 2003). Our lab demonstrated that 10μg/mL myricetin, and 10mg/mL $NaWO_4$, served as very strong inducers of the BaeSR pathway (Kurach, Oates, and Raivio unpublished). For the purpose of this experiment, we constructed strains CO88-CO92,

which contain the λp*lac*Mu53 insertions in *fecD,* near *yjgX,* in *mdtA, yggC,* and *yliA,* respectively, in a Δ*baeR* background. This was to allow us to determine if any observed induction in a wildtype strain background was in fact dependent on the activity of the BaeSR pathway. We compared the phenotypes of these new strains to those of strains CO23, CO16, CO4, CO2, and CO8, containing the insertions in an otherwise wildtype background, on MacConkeys Media to determine the effect of deletion of BaeSR on the basal expression of these genes. The results of these comparisons are depicted in Figures 3.8 and 3.9. We observed that the expression of the *lacZ* reporter in *yggC* appeared to be decreased in the absence of BaeR, as would be expected if BaeR was functioning as an activator of *yggC*. Likewise, the λp*lac*Mu53 insertions in *mdtA* and *yliA* also appeared to be less active or of similarly low expression level as the basal level of expression for these insertions. When we observed the λp*lac*Mu53 insertions in *fecD* and near *yjgX*, we saw that the expression of these *lacZ* reporters appeared to be increased in the absence of BaeR, as we would expect if BaeR was indeed acting as a repressor of these loci.

We used β-galactosidase assays to examine induction of β-galactosidase production by strains carrying the insertion mutations in response to known activators of the BaeSR pathway. Assays were done in the presence and absence of 2mM indole, 10μg/mL myricetin, and 10mg/mL $NaWO_4$, in exponential and stationary phase for both the original λp*lac*Mu53 insertion strains, and the *baeR* null strains derived from them. For all of the inducing cues used, under all conditions examined, we were unable to observe induction of β-galactosidase production in strains carrying the λp*lac*Mu53 insertions located in *yggC,* and *yliA,* or repression of β-galactosidase production in strains carrying

**Figure 3.8. λp*lac*Mu53 insertions in *mdtA, yggC and yliA* in a Δ*baeR* background**

P1*vir* lysates were prepared from strains hosting λp*lac*Mu53 insertions identified in our screen as being upregulated in the presence of BaeR overexpression and used to transduce the insertions into MC4100 Δ*baeR*. Single isolated colonies from each strain were then streaked out on Mac to observe the phenotype. The expression of β-galactosidase from strains carrying λp*lac*Mu53 insertions in *mdtA, yggC,* and *yliA,* strains CO4 and CO90, CO2 and CO91, and CO8 and CO92 respectively, is decreased or unaffected in a Δ*baeR* background, strains CO90, CO91, and CO92. Note that innate expression of *mdtA* and *yliA* is quite low.

A) CO4, λp*lac*Mu53 in *mdtA,* MC4100 background B) CO90, λp*lac*Mu53 in *mdtA,* Δ*baeR* background C) CO2, λp*lac*Mu53 in *yggC,* MC4100 background D) CO91, λp*lac*Mu53 in *yggC,* Δ*baeR* background E) CO8 λp*lac*Mu53 in *yliA,* MC4100 background F) CO92 λp*lac*Mu53 in *yliA* Δ*baeR* background.

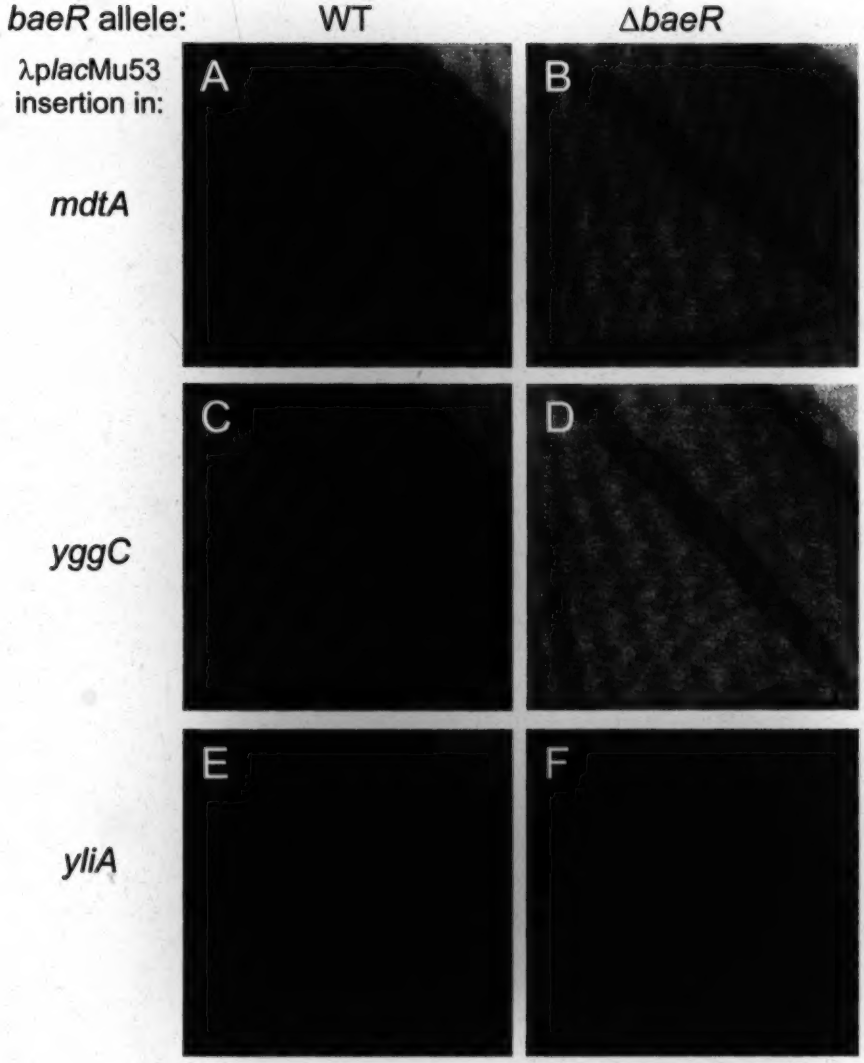
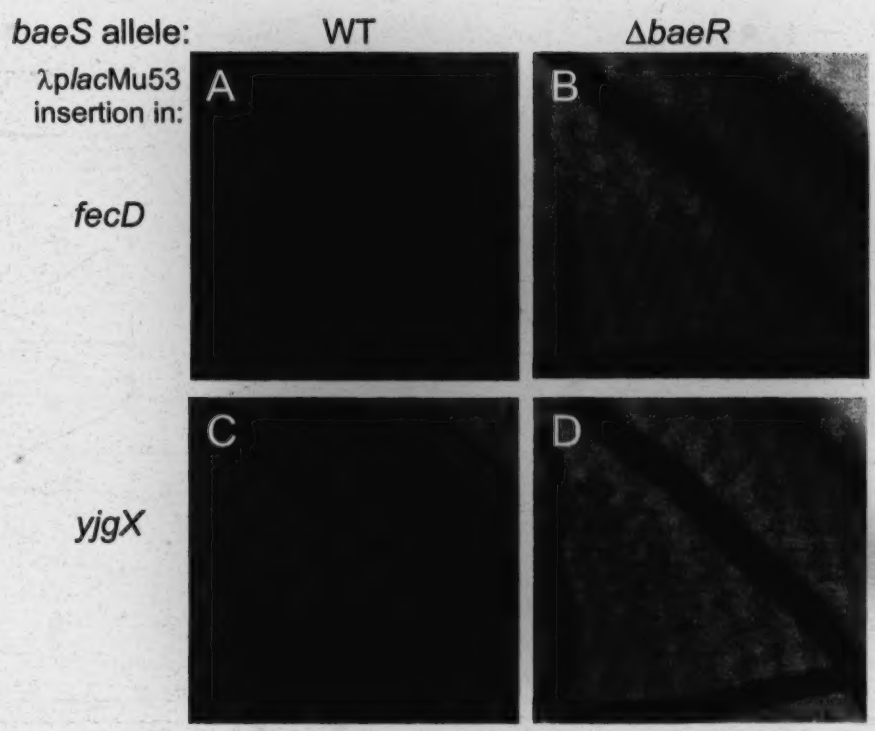

baeR allele:
WT
ΔbaeR
λplacMu53 insertion in:
mdtA
yggC
yliA
A
B
C
D
E
F

**Figure 3.9. λp*lac*Mu53 insertions in *fecD* and near *yjgX* in a Δ*baeR* background**

P1*vir* lysates were prepared from strains hosting λp*lac*Mu53 insertions identified in our screen as being downregulated in the presence of BaeR overexpression and used to transduce the insertions into MC4100 Δ*baeR*. SICs from each strain were then streaked out on Mac to observe the phenotype. The expression of β-galactosidase form strain carrying λp*lac*Mu53 insertions in *fecD* and near *yjgX*, strains CO23 and CO88, and CO16 and CO89 respectively, appeared to be increased in a Δ*baeR* background. Note that *yjgX* appears to be expressed in the presence of wildtype levels of BaeR as well.

A) CO23, λp*lac*Mu53 in *fecD*, MC4100 background B) CO88, λp*lac*Mu53 in *fecD*, Δ*baeR* background C) CO16, λp*lac*Mu53 near *yjgX*, MC4100 background D) CO89, λp*lac*Mu53 near *yjgX*, Δ*baeR* background.


baeS allele:
WT
ΔbaeR
λplacMu53 insertion in:
fecD
yjgX
A
B
C
D

the λp*lac*Mu53 insertions located in *fecD,* and near *yjgX.* We did observe induction of β-galactosidase expression in the strain bearing a λp*lac*Mu53 insertion located in known BaeSR regulon member *mdtA* (Figure 3.10). We found that the presence of 2mM indole resulted in a 2.6 fold increase in the expression of *mdtA-lacZ* in cultures in exponential phase, a value in line with those previously reported, which ranged from 2.4 to 3.9 fold (Nishino *et al.* 2005, Hirakawa *et al.* 2005). Strikingly, the increases in expression resulting from the novel inducing cues myricetin and $NaWO_4$ were much higher than this, being 6.9 and 11.6 fold respectively. This suggests that the novel inducing cues identified by our group are able to activate the BaeSR pathway to a greater extent than indole, which had previously been used as a general inducer of this pathway. For all three inducing cues, this induction was eliminated in the absence of BaeR resulting in fold induction changes of 1.0 by indole, 1.3 by myricetin and 1.2 $NaWO_4$. This suggests that induction in response to these cues is primarily mediated via the BaeSR pathway. For indole, this observation corresponds to the absence, or substantial reduction, of *mdtA* induction that has been previously reported in a *baeR* null background (Nishino *et al.* 2005, Hirakawa *et al.* 2005). Finally, a comparison of *mdtA-lacZ* expression in wildtype and Δ*baeR* backgrounds (strains CO4 and CO90) in the absence of induction confirmed that there was a slight reduction of *mdtA* expression in the absence of BaeR, corresponding to a 0.8 fold decrease. None of the inducing cues we examined were able to induce the expression of *mdtA* in a stationary phase culture.

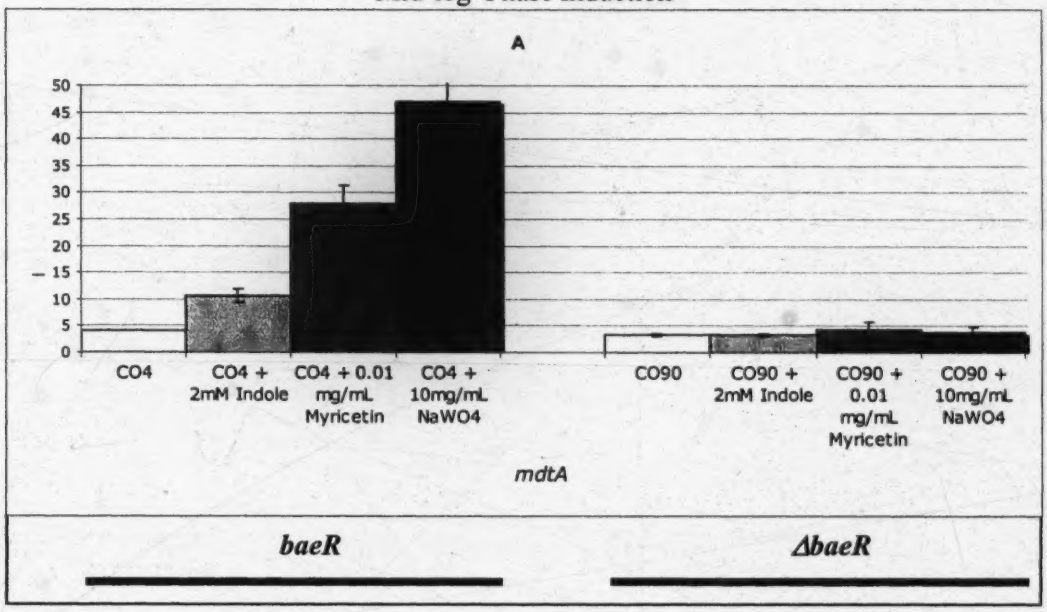
**Figure 3.10. Induction of β–galactosidase in strains carrying a λp*lac*Mu53 insertion in *mdtA* by inducers of the BaeSR pathway.**

Five samples of each strain were grown overnight, then subcultured 1:50 and allowed to grow for three hours at 37˚C. Induction was carried out by the addition of 2mM indole, 10μg/ml myricetin, or 10mg/mL $NaWO_4$, and incubation was continued for two hours. At this point a β–galactosidase assay was performed as described in Slauch and Silhavy (1991). The $OD_{600}$ and $A_{420}$ were determined and converted into β–galactosidase units using the formula $MU=(\Delta A_{420}/time * 1000 * 60)/ OD_{600}$. These units appear along the Y axis of each graph. For each strain, two outlying values were discarded, and the remaining data points were averaged and the standard deviation was determined. Data are provided for three hour subcultures and overnight cultures. This experiment was replicated with similar results (not shown).

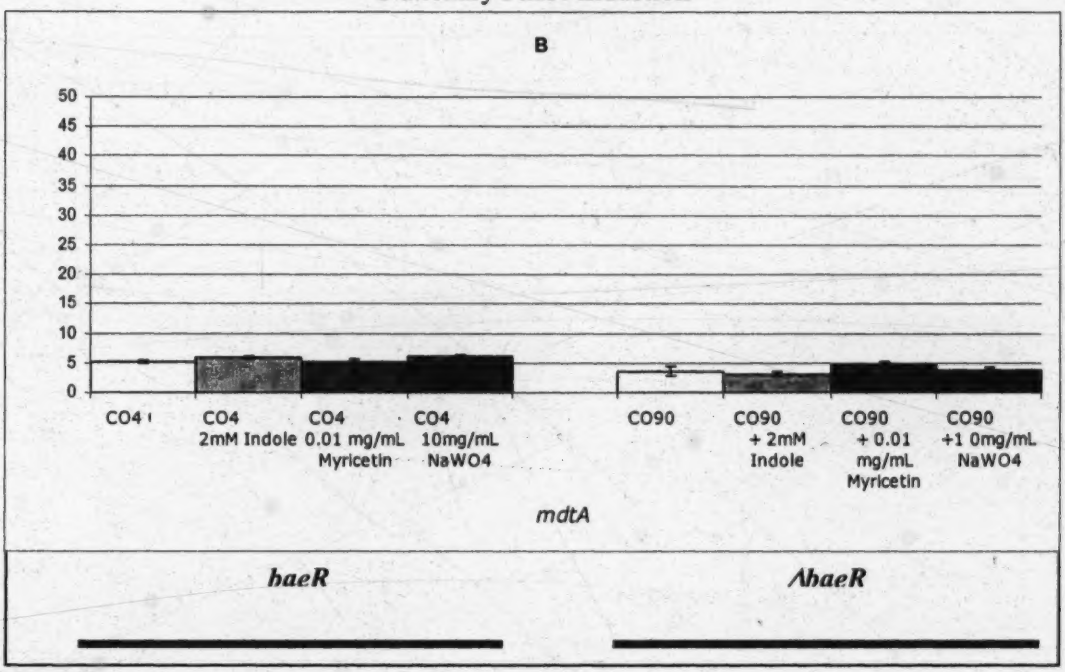
A) β–galactosidase production in 3 hour subcultures (Mid-log phase induction) of strains containing a λp*lac*Mu53 insertion in *mdtA* in a WT (left) or Δ*baeR* (right) background induced with 2mM indole, 10μg/mL myricetin, or 10mg/mL $NaWO_4$ as indicated.

B) β–galactosidase production in O/N cultures (Stationary phase induction) of strains containing a λp*lac*Mu53 insertion in *mdtA* in a WT (left) or Δ*baeR* (right) background induced with 2mM indole, 10μg/mL myricetin, or 10mg/mL $NaWO_4$ as indicated.

Mid-log Phase Induction




Stationary Phase Induction



# 4. Molecular Characterization of the BaeSR regulon

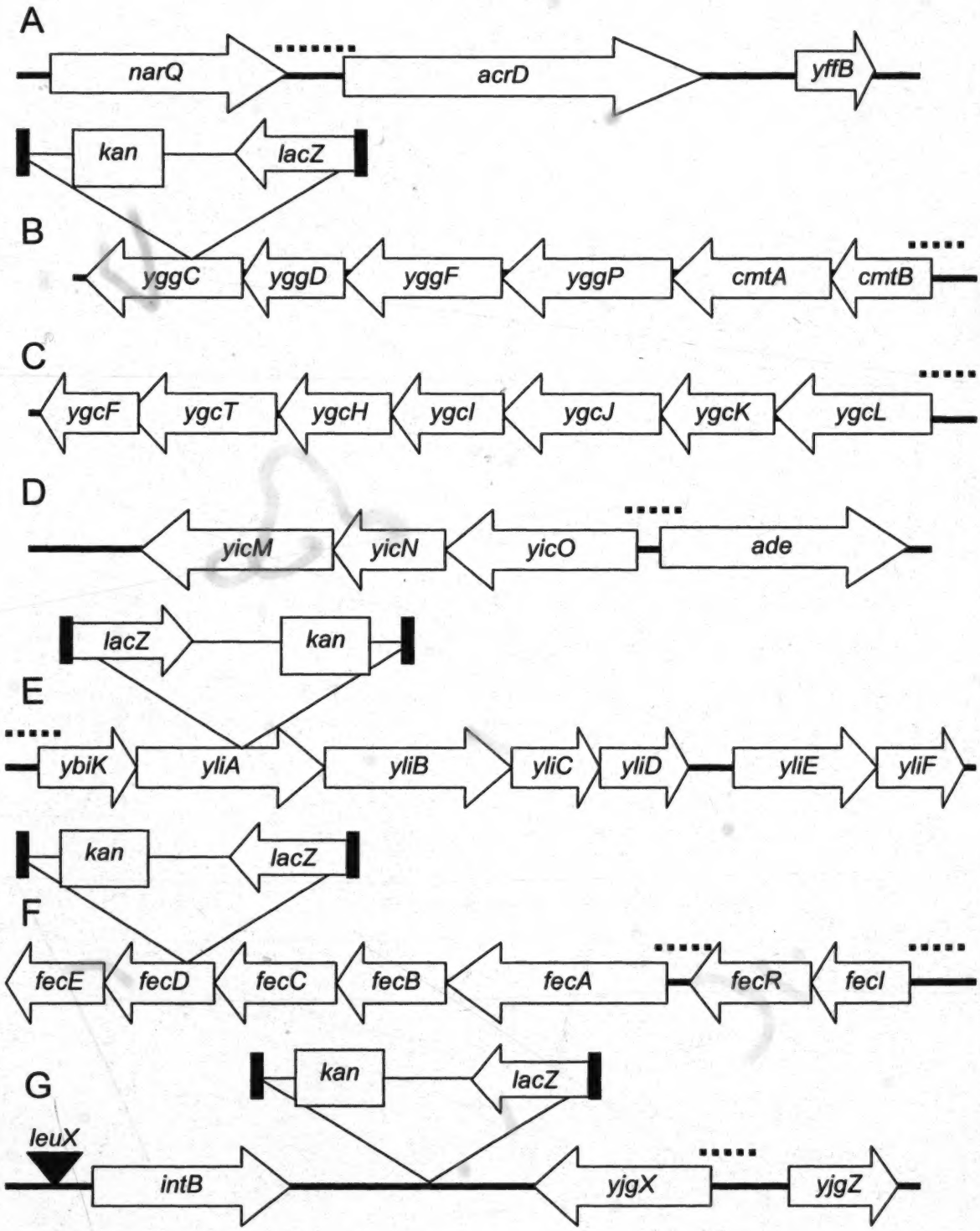
## 4.1 BaeR binding to promoter regions of putative regulon members

In order to determine if the genes into which we identified insertions in our screen were directly controlled by the BaeSR proteins, we performed electrophoretic mobility shift assays (EMSAs) to determine if a MBP-BaeR fusion protein was able to bind to their promoter regions. As many of the λp*lac*Mu53 insertions appeared to be located in genes which were members of operons based on their chromosomal contex, we selected the intergenic regions upstream of the first putative operon member for study. The locations of the fragments selected for this study relative to the locations of the λp*lac*Mu53 insertions are indicated in Figure 4.1. In addition to these regions, we looked at the promoter region for *fecIR* as these genes represent characterized regulators of the *fecABCDE* operon. Furthermore, we looked at the promoter regions of two genes, *yicO* and *ygcL* which were not identified in this study, but were identified in the literature as being activated by BaeR (Baranova and Nikaido 2002). The promoter region of AcrD was also assayed for MBP-BaeR binding, as this canonical regulon member could serve as a positive control. The locations of the fragments selected to study BaeR binding to this second set of promoters are also indicated in Figure 4.1.

**Figure 4.1. Locations chosen for EMSA study**

These diagrams depict the locations and directions of the operons that were examined by electrophoretic mobility shift assay. We chose to examine the intergenic regions upstream of the operons that our λp*lac*Mu53 insertions were located in for BaeR binding. In addition to this, we also examined the promoter regions of the *acrD, ygcL,* and *yicO* genes, which had previously been reported to be BaeSR regulated, and the promoter region of the *fecIR* genes, which are known to regulate *fecABCD*. The locations of the PCR fragments utilized in this study are indicated by the dashed lines.

A) *acrD* B) λp*lac*Mu53 insertion in *yggC*. C) *yicONM* operon D) *ygcLKJIHTF* operon E) λp*lac*Mu53 insertion in *yliA*. F) λp*lac*Mu53 insertion in *fecD*. G) λp*lac*Mu53 insertion near *yjgX*.


A
narQ
acrD
yffB
kan
lacZ
B
yggC
yggD
yggF
yggP
cmtA
cmtB
C
ygcF
ygcT
ygcH
ygcI
ygcJ
ygcK
ygcL
D
yicM
yicN
yicO
ade
lacZ
kan
E
ybiK
yliA
yliB
yliC
yliD
yliE
yliF
kan
lacZ
F
fecE
fecD
fecC
fecB
fecA
fecR
fecI
kan
lacZ
G
leuX
intB
yjgX
yjgZ

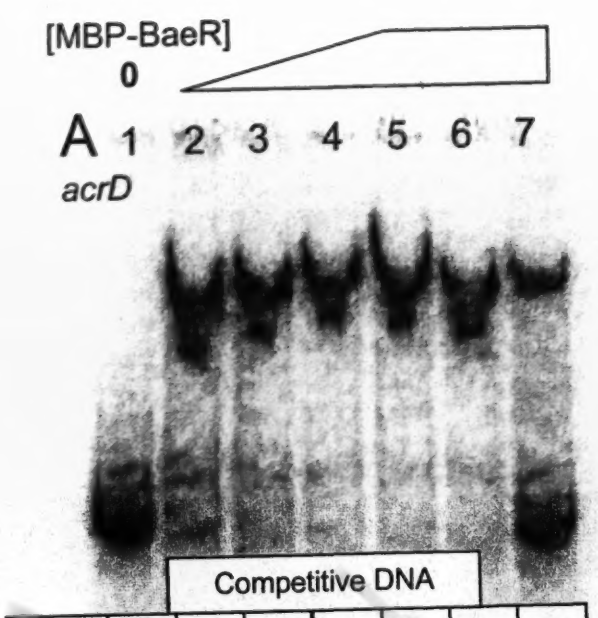
Briefly, our EMSA studies utilized an affinity purified sample of MBP-BaeR and $^{32}$P-labled PCR fragments from the promoter regions under study. As response regulators are activated as transcriptional regulators *in vivo* upon phosphorylation, we utilized a period of incubation of MBP-BaeR with the small phosphor-donor acetyl-phosphate in order to simulate this physiological situation. In order to ensure that any binding observed by these studies represented specific binding of MBP-BaeR to the promoter fragments under study, we utilized 50μg/mL poly(dI-dC) in each reaction, as well as an excess of unlabeled *cpxP* promoter, which we had previously observed not to be bound by MBP-BaeR, as a non-specific competitor in a redundant reaction containing the highest amount of protein assayed. Furthermore, we used an excess of an unlabeled sample of the promoter under study to titrate MBP-BaeR away from the labeled fragment, again demonstrating that the binding we observed was specific for the fragment under study. The results of our EMSA studies for promoter fragments that did bind MBP-BaeR~P are depicted in Figure 4.2. We used a fragment of the *acrD* promoter as a positive control for binding as this promoter had previously been shown to be bound by BaeR (Hirakawa *et al.* 2005). We found that all of the *acrD* promoter was shifted at the lowest concentration of MBP-BaeR we used, 0.8μM (Figure 4.2A, lane 1). This experiment demonstrated that the MBP-BaeR protein we utilized was able to recognize and bind DNA at concentrations similar to those reported previously in the literature, which ranged from 16nM to 2.3μM, with most values being above 0.5μM (Baranova and Nikaido 2002, Hirakawa *et al.* 2005). It also demonstrated the effectiveness of our specificity controls, as the presence of unlabeled *cpxP* promoter did not affect MBP-BaeR binding to the *acrD* promoter,

**Figure 4.2. Electrophoretic mobility shift assays demonstrating MBP-BaeR binding to the *acrD*, *ygcL*, *yicO*, and *cmtB* promoter regions.**

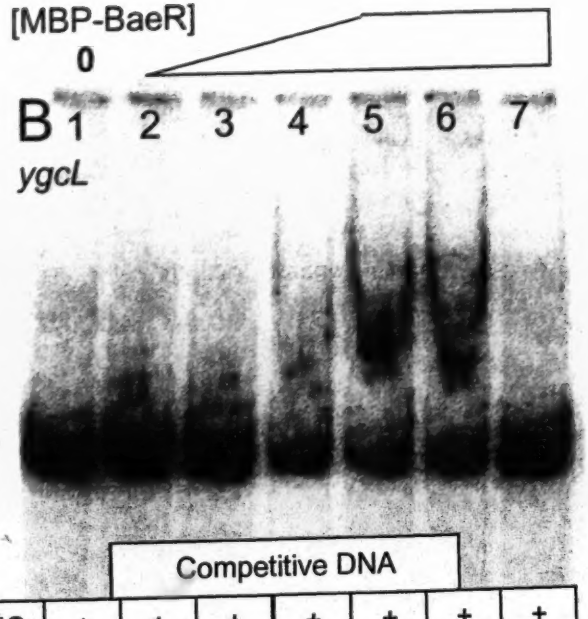
Assays were carried out using a purified radio-labeled sample of the promoter region indicated incubated with varying concentrations of MBP-BaeR~P in the presence of 10mM Tris-Cl (pH 7.5), 50mM KCl, 1mM dithiotheritol, 2.5% glycerol, 0.05% NP-40, and 50ng/μL poly(dI-dC). These reactions were loaded on a polyacrylamide slab gel which was run for 1 hour at 100V, then dried and visualized by overnight exposure to a phospho-imaging screen.

A) MBP-BaeR binding to the *acrD* promoter, B) MBP-BaeR binding to the *ygcL* promoter, C) MBP-BaeR binding to the *yicO* promoter, D) MBP-BaeR binding to the *cmtB* promoter.

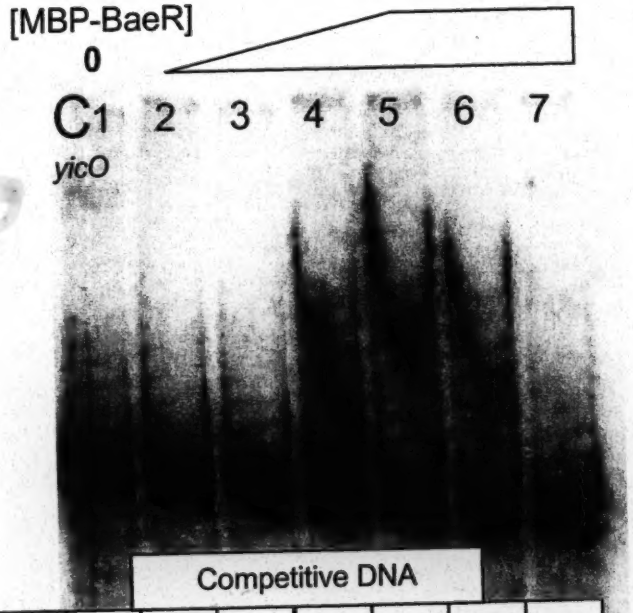
Lane identities are as follows for each image: 1: No protein, 2: 0.8μM MBP-BaeR, 3: 1.0μM MBP-BaeR, 4: 1.5μM MBP-BaeR, 5: 2.0μM MBP-BaeR 6: 2.0μM MBP-BaeR + excess unlabeled *cpxP* as a non-specific competitor, 7: 2.0μM MBP-BaeR + excess unlabeled promoter as a specific competitor.



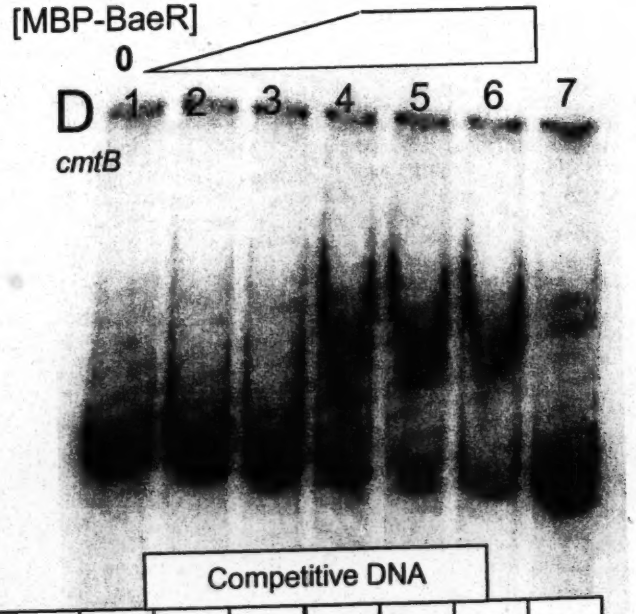

| Competitive DNA | | | | | | | |
|---|---|---|---|---|---|---|---|
| dI-dC | + | + | + | + | + | + | + |
| *cpxP* | - | - | - | - | - | + | - |
| *acrD* | - | - | - | - | - | - | + |




| Competitive DNA | | | | | | | |
|---|---|---|---|---|---|---|---|
| dI-dC | + | + | + | + | + | + | + |
| *cpxP* | - | - | - | - | - | + | - |
| *ygcL* | - | - | - | - | - | - | + |




| Competitive DNA | | | | | | | |
|---|---|---|---|---|---|---|---|
| dI-dC | + | + | + | + | + | + | + |
| *cpxP* | - | - | - | - | - | + | - |
| *yicO* | - | - | - | - | - | - | + |




| Competitive DNA | | | | | | | |
|---|---|---|---|---|---|---|---|
| dI-dC | + | + | + | + | + | + | + |
| *cpxP* | - | - | - | - | - | + | - |
| *cmtB* | - | - | - | - | - | - | + |

while unlabeled *acrD* promoter reduced the shifted band considerably (Figure 4.2A Lanes 6 and 7). We also observed specific MBP-BaeR binding to the promoter regions of *ygcL, yicO*, and *cmtB*, although only at higher protein concentrations than were necessary for *acrD* binding. For the *ygcL* we observed binding beginning in the presence of 1.5μM MBP-BaeR (Figure 4.2B Lanes 4 and 5). This binding was not affected by the presence of unlabeled *cpxP* promoter, however excess unlabeled *ygcL* promoter eliminated it entirely (Figure 4.2B Lanes 6 and 7). We observed similar results with the *yicO* promoter, with binding beginning in the presence of 1.5μM MBP-BaeR (Figure 4.2C Lanes 4 and 5). The appearance of the unbound promoter appearing slightly darker in the presence of *cpxP* promoter may be considered a pippeting error resulting in a larger quantity of labled probe for two reasons, firstly the shifted complex is unaltered in intensity compared to lane five, and secondly as work done in our lab has indicated that the cpxP promoter fragment is not able to be bound by MBP-BaeR. Excess unlabeled *yicO* promoter eliminated binding by MBP-BaeR to the labled fragment entirely (Figure 4.2C Lanes 6 and 7). Finally, we observed binding to the *cmtB* promoter region over MBP-BaeR concentration range of 1.5-2.0μM (Figure 4.2D Lanes 4 and 5). This binding was specific, as it was not affected by the *cpxP* promoter, while the shifted complex was not seen in the presence of excess unlabeled *cmtB* promoter (Figure 4.2D Lanes 6 and 7). These observations are consistent with the lower levels of induction in response to BaeR overexpression observed for *yicO* and *ygcL* when compared with canonical BaeSR regulon member *mdtA,* and with our inability to observe induction of *cmtB* in response to the BaeSR inducing cues indole, myricetin and $NaWO_4$ (Baranova and Nikaido 2002,

Section 3.3). We did not observe any binding to the promoters of *ybiK, fecA, fecI,* or *yjgX.*

**4.2 Identification of putative BaeR binding recognition sequences**

Since a consensus binding sequence has been proposed for BaeR, we conducted a visual search of the upstream regions of each of the genes for which we observed MBP-BaeR binding, as well as those genes identified in our screen which did not bind BaeR, for sequences matching the consensus sequence (Nishino *et al.* 2005). Since the consensus sequence proposed for BaeR is T rich, we searched the 250bp upstream of the start codon of the indicated genes for T rich regions by visual inspection. The sequences identified in this manner were then visually examined in greater detail in order to detect the possible presence of sequences with similarity to the consensus sequence. As a sequence in the *ycaC* promoter showing 50% identity to the consensus sequence has previously been identified as a BaeR recognition sequence, only sequences showing greater than 50% identity were identified in this study (Nishino *et al.* 2005). Using this method we located sequences with similarity to the consensus sequence in the promoter regions of *yicO, ygcL, cmtB,* and *fecI* (Table 4.1). As *acrD* was one of the regulon members used in the original study we followed the suggestion of Nishino *et al.* in identifying the recognition sequence in the *acrD* promoter. The sequences we identified have 61% identity with the consensus sequence for *ygcL, cmtB,* and *fecI,* while an area of 67% identity was located in the *yicO* promoter. Another putative regulon member, *ycaC*

**Table 4.1 Putative BaeR recognition sequences**

| Gene Name | Recognition Sequence | Location† | Percent Identity |
|---|---|---|---|
| Consensus | TTTTTCTCCATD*ATTGGC | | |
| *acrD* | TTTTTCTCCACGATTGGC | -88 to -71 | 94% |
| *yicO* | TTTTTCTC-ATTTTGATT | -95 to -77 | 67% |
| *cmtB* | TTTCCCTGCTCAATTGCT | -88 to -70 | 61% |
| *ygcL* | TTTTTACCGATAACGGAA | -73 to -55 | 61% |
| *fecI* | TTTTATTTCC-AATTGTA | -76 to -59 | 61% |
| *ycaC* | TATTTCCCGTCTATGCTT | -82 to -65 | 50% |

The consensus sequence was initially proposed based on computational analysis of the *acrD, mdtA* and *spy* promoters. The recognition sequences for *acrD* and *ycaA* are those indicated in this study (Nishino *et al.* 2005).

* where D is G, T, or A

†Location is given relative to the start codon of the gene

Underlined basepairs indicate conformity to the consensus sequence

has been noted to contain an area of only 50% conformity to the consensus sequence (Nishino *et al.* 2005). The extent of conformity to the consensus sequence observed in the well established BaeSR regulon members *mdtA, spy,* and *acrD* is quite high, being 94%, 100%, and 94% respectively. A parallel can be noted between regulon members with promoter sequences containing an area of high homology to the consensus sequence, and putative regulon member promoter sequences bearing lower homology to it, in that those members with the highest homology to the consensus sequence have been reported to be inducible to a higher extent. All of the regions of homology to the consensus sequence we identified fell within an area from -55bp to -95bp upstream of the start codon of the gene. This is comparable to the locations of the recognition sequences located in the promoters of the canonical BaeR regulon members, which are located between -55bp to -156bp from the start codon (Nishino *et al.* 2005). Since we did not observe any BaeR binding to the *fecI* promoter fragment tested in section 4.1, it is possible that this promoter requires the binding of a second regulator in order for BaeR binding to occur.

### 4.3 DNaseI protection assay of the *acrD* promoter

In order to further characterize the binding sequence of MBP-BaeR, we undertook DNaseI protection assays to identify the sequence recognized by this protein. For our initial experiments we opted to use the *acrD* promoter region, as this gene had recently been shown to belong to the BaeSR regulon, and earlier electrophoretic mobility shift assays (Section 4.1) had demonstrated BaeR binding to this region (Hirakawa *et al.* 2003a, Hirakawa *et al.* 2003b). Using a PCR product radio-labeled on the *acrD* coding strand, we were able to obtain a clear region of protection from DNaseI degradation that

increased at elevated concentrations of MBP-BaeR (Figure 4.3). By comparing this region of protection with an adjacent sequencing ladder, we determined that the sequence bound by MBP-BaeR was 5'–GATAATTTACATTAACTCCTTTTTTTCTCCACGAT-3', representing position 2585506-2585540bp in the genome, and -71 to -109 from the *acrD* start codon. Shortly after this sequence was obtained, a study was published containing an experimentally determined footprint for BaeR binding to the *acrD* promoter. The sequence bound was identified as being a 32bp region -71 to -105bp from the *acrD* start codon (Hirakawa *et al.* 2005). This corresponds to our sequence with a great degree of precision, differing only by four bp, so taken together, these two results corroborate each other. Sequence alignment of the promoter regions of known regulon members *spy, acrD, mdtABCD* and the putative regulon member *ycaC* was used by another research group to propose a consensus sequence for BaeR binding (Nishino *et al.* 2005). The recognition sequence they proposed for BaeR binding to the *acrD* promoter differed from the one experimentally verified above by four basepairs being from -88 to -71 from the *acrD* start codon, Figure 4.4. In contrast, for the only other regulon member for which we have DNaseI protection assay data , *mdtA,* their proposed recognition binding site overlapped the experimentally determined binding site for only 5 bp out of 18 total (Hirakawa *et al.* 2005, Nishino *et al.* 2005). While the computationally determined consensus sequence proposed in Nishino *et al.* 2005 is highly conserved amongst the established regulon members, and it is near the binding sites established for *acrD* and *mdtA,* DNaseI protection assay data should be obtained for *spy*

**Figure 4.3. DNaseI protection assay of the *acrD* promoter coding strand.**

This DNaseI protection assay was carried out using a radio-labeled section of the *acrD* promoter incubated with varying amounts of MBP-BaeR in a buffer containing 5mM MgCl, 12.5% glycerol, 10mM Tris-Cl (pH 7.5), 50mM KCl, 1mM dithiotheritol, and 0.05% NP-40. These samples were subjected to digestion for 30sec by the addition of 0.025U DNaseI followed by phenol:chloroform extraction. The products were precipitated, resuspended in 4μL formamide loading dye, and run on a sequencing gel adjacent to a sequencing ladder at 40W for 1.5 hours. The gel was dried and visualized by overnight exposure to a phospho-storage screen.

Lane identities are as follows: 1: 1.5μM MBP-BaeR, 2: 1.2μM MBP-BaeR, 3: 0.9μM MBP-BaeR, 4: 0.6μM MBP-BaeR, and 5: 0.0μM MBP-BaeR. Lanes C, T, A, and G represent a sequencing ladder. The thick black bar indicates the area of the *acrD* promoter protected by the increasing MBP-BaeR concentration.

Based on comparison with the sequencing ladder, the site protected by MBP-BaeR binding to the *acrD* promoter is:

5'-GATAATTTACATTAACTCCTTTTTTTCTCCACGAT-3'

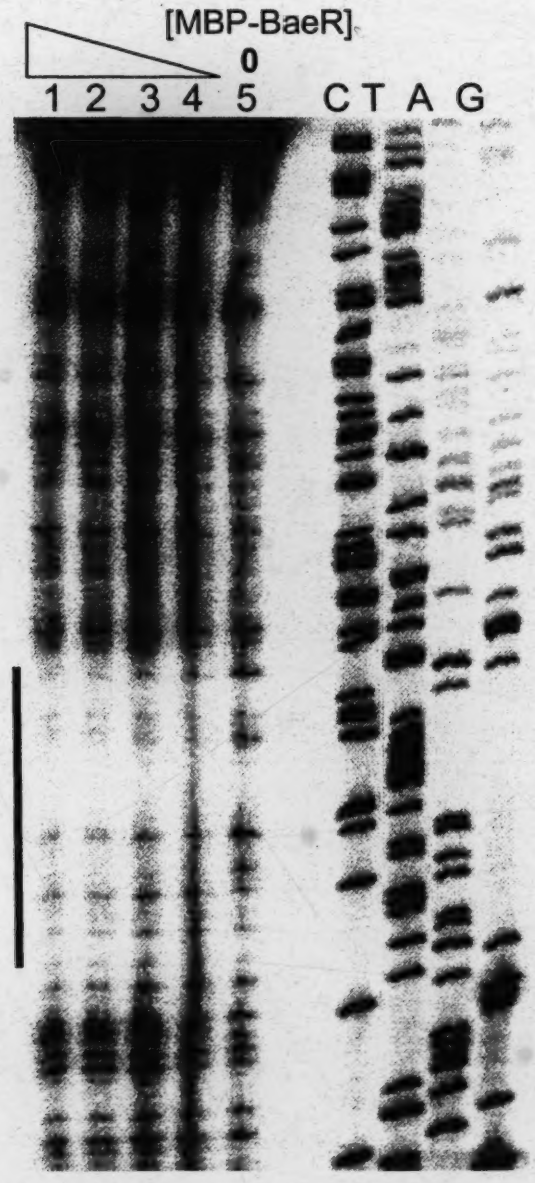
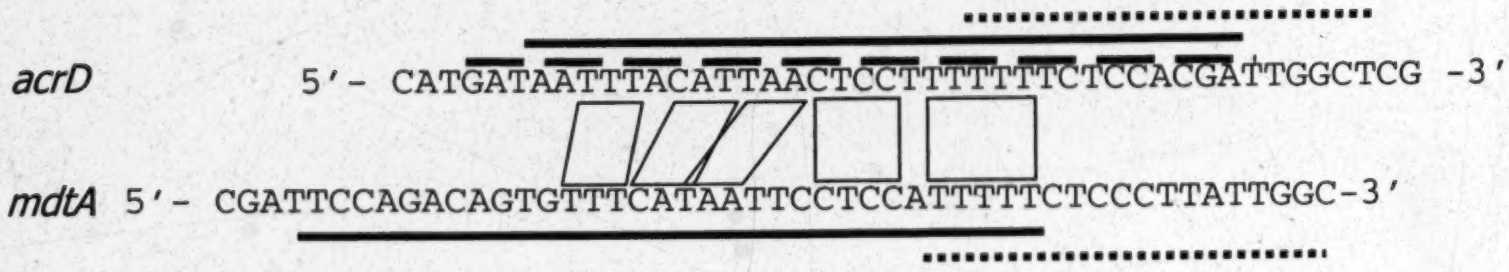

[MBP-BaeR]
0
1 2 3 4 5
C T A G

**Figure 4.4. Alignment of the BaeR binding sites in the *acrD* and *mdtA* promoters**

Depicted are portions of the *acrD* and *mdtA* promoter regions which contain the experimentally determined BaeR binding sites and the computationally determined recognition sequences. The solid black lines represent the BaeR binding sequences determined by Nishino *et al.* 2005. The dashed black line is the region of BaeR binding in the *acrD* promoter determined in this study. The dotted lines are the recognition sequences proposed by the computational study, demonstrating only partial overlap with the experimentally determined binding sequences. The clear boxes indicate regions of similarity between the binding sites of BaeR in the *acrD* and *mdtA* promoters. With additional data from the *spy* promoter, this may serve to produce a new consensus binding sequence.


acrD 5′- CATGATAATTTACATTAACTCCTTTTTTTCTCCACGATTGGCTCG -3′
mdtA 5′- CGATTCCAGACAGTGTTTCATAATTCCTCCATTTTTCTCCCTTATTGGC-3′

as well to establish the validity of this sequence. As is demonstrated in Figure 4.4, the BaeR binding sequences experimentally determined by DNaseI protection assay in the *mdtA* and *acrD* promoters exhibit significant sequence similarity. It is possible that with the addition of a BaeR binding sequence for the *spy* promoter this similarity could be utilized to identify a new BaeR binding consensus, which could be expected to conform better to experimentally determined BaeR binding than the computationally produced consensus. DNaseI protection assays should be performed on the promoters of *yicO, ygcL,* and *cmtB* to verify the putative regions of conformity to the consensus sequence in these promoters.

# 5. Discussion

Prior to this study, the BaeSR system had been implicated in the envelope stress response due to its co-regulation of the stress protein Spy with the CpxRA envelope stress response pathway, and due to its response to the envelope stresses spheroplasting, PapG overexpression, and indole (Raffa and Raivio 2002). BaeSR had also been shown to be involved in the regulation of multidrug efflux, regulating the expression of two pumps, MdtABCD, and AcrD, which confer resistance to novobiocin, β-lactams, and bile salts (Baranova and Nikaido 2002, Nagakubo *et al.* 2002, Hirakawa *et al.* 2003). Since it had been noted that a *baeR*, *cpxR* double null strain had increased sensitivity to PapG overexpression compared to either single mutant, and that this effect was not due to the Spy protein, it was suggested that BaeSR may control more genes involved in the envelope stress response (Raffa and Raivio 2002). This thesis described the use of a λp*lac*Mu53 transposon screen to identify novel members of the BaeSR regulon and to better characterize its physiological role.

## 5.1 Diversity of the BaeSR regulon

### 5.1.1 Novel genes identified in this study

This project utilized a λp*lac*Mu53 transposon screen to identify genes which responded to the overexpression of BaeR. We identified a total of five genes which appear to be BaeR regulated, responding to both BaeR overexpression, and the presence of *baeS1*::Tn*10cam*, a *baeS* gain of function allele which constitutively activates the

BaeSR pathway. Four of the genes which we identified, *yliA*, *yjgX*, *yggC*, and *fecD* represent new BaeSR regulon members, while *mdtA* has been previously characterized as being in the BaeSR regulon. We found that *mdtA*, *yliA*, and *yggC* are activated by the pathway, while *yjgX* and *fecD* appear to be repressed by it.

#### 5.1.1.1 *yliA*

This thesis presented evidence that expression of the *yliA* gene was upregulated by BaeR overexpression and the *baeS1*::Tn*10cam* allele (Figures 3.2, 3.5, and 3.7). The *yliA* gene is believed to be part of an operon consisting of *ybiKyliABCD*. The function of *ybiK* was suggested to be related to the transport of glutathione however it was later demonstrated to be dispensable for this function (Parry and Clark 2002, Suzuki *et al.* 2005). The *yliABCD* genes were found to encode the actual glutathione uptake system, representing the first such system identified in bacteria (Suzuki *et al.* 2005). The *yliABCD* glutathione uptake system was found to be only partially responsible for the uptake of glutathione, with a more substantial role played by the γ-glutamyltranspeptidase GGT. A near total reduction of glutathione in the media was found to result from overexpression of *yliABCD*, while the deletion of *yliABCD* resulted in only a slight reduction in the rate glutathione was removed from the media (Suzuki *et al.* 2005). This suggests that the expression of *yliA* is generally very low, as we noted for CO8 (*yliA*::λp*lac*Mu53), compared to CO59 (*yliA*::λp*lac*Mu53, *baeS1*::Tn*10cam*) (Figure 3.7). Furthermore, glutathione concentration is known to decrease in stationary phase cultures, and *ybiK* is known to be less expressed in rich media, confirming our speculation of growth phase dependent regulation with regards to the increase in expression observed for strain CO59

(*yliA::*λ*plac*Mu53, *baeS1::*Tn*10cam*) in stationary phase (Suzuki *et al.* 2005, Parry and Clark 2002, Figure 3.7). Glutathione is abundant in aerobically growing cells, and is thought to be involved in the oxidative stress response and the detoxification of drugs which are electrophiles (Fahey *et al.* 1978, Becker-Hapak and Eisenstark, 1995, Peninckx and Elskins 1993). While glutathione appears to not be necessary for the oxidative stress response, the activity of the oxidative stress response protein OxyR is increased in its absence, suggesting that it does play some role (Greenberg and Demple 1986, Prieto-Álamo *et al.* 2000). The observation that glutathione is involved in oxidative stress response is of great interest to us as the BaeSR inducing cue indole is believed to disrupt the membrane in a manner leading to oxidative stress (Garbe *et al.* 2000). Therefore our identification of the glutathione uptake system YliABCD as being BaeSR regulated suggests a role for the BaeSR pathway in alleviating oxidative stress.

Another intriguing possibility is raised by the observation that *ybiK* is a member of the CysB regulon (Parry and Clark, 2002). CysB is a global regulator of genes involved in the biosynthesis and transport of L-cysteine, sulphate and thiosulphate transport and sulphate reduction (Kredich 1996). As we did not observe BaeR binding to the promoter of *ybiK* it is possible that BaeR regulates a regulator of this system making *cysB* a possible candidate. Mutations in CysB have been found to have increased resistance to novobiocin, however the mechanism for this has not been determined (Rakonjac *et al.* 1991). Work done in our lab has shown that CO8 (*yliA::*λ*plac*Mu53) also exhibits increased resistance to novobiocin suggesting that the resistance that occurs upon CysB mutation is a result of decreased expression of *yliABCD* (Kurach and Raivio unpublished). It is possible that as a transporter complex, YliABCD is leaky, permitting

the influx of novobiocin across the cellular membrane. Therefore, our results indicate that BaeSR upregulates both *yliABCD,* which would be expected to increase novobiocin sensitivity, and *mdtABCD,* which is known to increase novobiocin resistance (Baranova and Nikaido 2002). It is possible that in a natural environment, the BaeSR pathway elevates the expression of multidrug efflux pumps to counteract its elevation of the expression of leaky transporter proteins.

In our experiments we were unable to identify BaeR binding to the *ybiK* promoter, or observe a sequence with similarity to the BaeR consensus sequence in this region. It is possible that BaeR regulates *ybiKyliABCD* in a secondary manner by regulating a regulator of this pathway. A possibility for this factor would be CysB. A more attractive idea however may be suggested based on the observation that BaeR and CpxR bind in close proximity to each other in the promoters of both *mdtABCD* and *acrD*, resulting in co-regulation (Hirakawa *et al.* 2005). It is possible that BaeR and CysB both bind the *ybiK* promoter together, however BaeR binding is too weak to assay in the absence of CysB.

#### 5.1.1.2 *yjgX*

We identified *yjgX* as being the nearest gene to a λp*lac*Mu53 insertion located in an intergenic region which had a phenotype of decreased *lacZ* expression in the presence of both BaeR overexpression and *baeS1*::Tn*10cam.* As we were surprised at the localization of an insert with this phenotype in an intergenic region we considered the possibility that this may represent the location of an sRNA, however none of those identified to date are located in this area (Hershberg *et al.* 2003). *yjgX* is thought to be a

pseudogene in *E. coli* K12 due to the presence of an inframe stop codon in its open reading frame, however it appears to be intact in other *E. coli* strains. By homology, it appears to belong to the sulphatase enzymes, which are responsible for hydrolyzing sulphate esters. We have found BaeR upregulates *yliA,* which is also regulated by CysB, which is not maximally active in the absence of sulphur limitation (Kredich 1996). It is possible that *yjgX* downregulation by BaeR represents a restriction of the cellular usage of sulphur during times when it is limited, representing a possible metabolic role for the BaeSR pathway. Both *yjgX* and *yliABCD* appear to play a role in cellular utilization of sulphur, with *yjgX* potentially playing a metabolic role, and *yliABCD* transporting glutathione, which can serve as a sulphur source. It is possible that this observation suggests a role for the BaeSR pathway in sulphur metabolism. However, because of the unexpected localization of the insertion point and the absence of BaeR binding to the *yjgX* promoter region, more work must be done to characterize *yjgX* as a member of the BaeSR regulon (Section 5.3).

#### 5.1.1.3 *yggC*

We identified *yggC* as being upregulated by both BaeR overexpression, and the *baeS1*::Tn*10cam* allele. *yggC* appears to be in an operon containing *cmtBAyggPFDC.* While it has homology to numerous cellular kinases, nothing has been published regarding the function or regulation of this gene. YggP, YggF and YggD have homology to oxidoreductases, GlpX a Fructose-1,6-bisphosphatase, and the mannitol repressor MtlR respectively, however, again nothing has been published regarding them. *cmtBA* appear to comprise a phosphoenol pryuvate: carbohydrate phosphotransferase system

(PTS), and *cmtA* has 51% identity with *mtlA*, a PTS mannitol transporter (Sprenger 1993). Studies were carried out which indicated *cmtBA* did not complement mutations in *mtlA* in either low or high copy number, and only partial complementation was observed using a heterologous promoter at high copy number (Sprenger 1993). As a result of this, and the fact that this study did not detect a promoter upstream from them lead to the suggestion that they were pseudogenes. I would question this conclusion, particularly as subsequent computational promoter searches have indeed proposed a promoter for *cmtBA* (http://regulondb.ccg.unam.mx:80/index.html). Furthermore, the function of *cmtBA* was assumed to be in mannitol transport based on homology to *mtlA* and as a result, mannitol transport was used to assay for their expression. It is possible that the reported homology to *mtlA* reflects conservation of the structure of the PTS components themselves, as opposed to functional conservation of the substrate recognized. Due to the overall level of homology, it is possible that over expression of *cmtBA* may allow uptake of mannitol even if it is not a primary substrate. Another possibility is that the *cmtBA* promoter is weak, and requires the binding of BaeR under inducing conditions for expression to occur. Our identification of *yggC* upregulation upon BaeR overexpression and demonstration of BaeR binding to the *cmtB* promoter, which contains a putative BaeR consensus sequence, suggests that that the BaeSR regulon contains members which are involved in metabolic functions like PTS transporters or have possible regulatory functions such as a kinase. These observations show that the BaeSR regulon contains members that may not be directly involved in either envelope stress response or multidrug efflux. We note with interest however, the possibility that, as *cmtBA* encode

integral membrane proteins, their expression may serve to complement the functions of other membrane transporters whose functions have been disrupted by envelope stress.

#### 5.1.1.4 *fecD*

*fecD* is the inner membrane component of the ferric citrate transporter FecABCDE. This thesis presents evidence that *fecD* is downregulated by the BaeSR system. This would make BaeSR the third system known to regulate *fecABCDE,* as it is known to be controlled by both the FecI alternative σ factor, and the iron repressor Fur (Buchanan 2005). FecI is able to function as a σ factor upon receiving a signal from FecR, an integral membrane anti-sigma factor, which in turn is transduced from FecA, the OM transporter, indicating the binding of ferric citrate at the cell surface. We examined both the *fecIR* and *fecABCDE* promoter regions for BaeR binding but were unable to detect any. We did however detect an area with homology to the BaeR consensus sequence in the *fecI* promoter, though the identity of this sequence was weaker than the others we identified in this study (Table 4.1). It is possible that BaeR does indeed exert its negative effect on *fecABCDE* expression via downregulation of *fecIR*, however it may require the binding of a second regulator in order to successfully bind the promoter. This model directly parallels the previously reported observation that BaeR co-regulates both *mdtABCD* and *acrD* in conjunction with CpxR, allowing higher levels of expression when both proteins are present (Hirakawa *et al.* 2005). Both *fecIR* and *fecABCD* are known to be downregulated by the ferric uptake regulator Fur, which represses a wide assortment of genes involved in iron uptake, storage and metabolism (Buchanan 2005, Hankte 2001). Fur becomes active as a transcriptional regulator upon binding $Fe^{2+}$,

indicating that intracellular iron levels are high. It is possible that BaeR binding to the *fecIR* promoter is only possible if it occurs concurrently with Fur binding. This suggestion is attractive as iron is essential for most life forms, and rare in most environments, so it may be undesirable to turn off an iron uptake system in response to envelope stressors (as BaeSR activating signals are thought to be) unless iron is also plentiful.

Despite its essential nature, iron represents a hazard to the cell when it is abundant, as it may produce hydroxyl radicals via the Fenton reaction. As such Fur is regulated by the oxidative stress responses OxyR and SoxRS (Zheng *et al.* 1999). Perhaps downregulation by BaeR is also a response to oxidative stress or the effects it may have on the envelope in conditions of high iron. Making this effect contingent upon downregulation mediated by Fur would ensure that iron uptake is only inhibited when iron is plentiful and may be stressing the cell, instead of as a result of BaeR induction by other pathway inducers. This idea conforms to our inability to see repression of *fecD* by indole, myricetin or $NaWO_4$.

Interestingly, *fecA* has also been observed to be downregulated by the MarA protein, which is known to control a wide variety of genes involved in resistance to multiple drugs and environmental hazards (Barbosa and Levy 2000). It has been reported that deletion of the extracellular loops of FecA abolish its function as a transport protein, while deletion of the extracellular loops of the related ferrichrome transporter FhuA convert it into an open channel and render the cell sensitive to antibiotics (Sauter and Braun 2004, Killmann *et al.* 1996). This demonstrates that these extracellular loops are not of conserved function, with those of FecA being involved in substrate uptake altering

their conformation upon ferric citrate uptake to block egress from the channel, while those of FhuA are involved in channel gating. These observations raise the possibility that the FecA transporter is somewhat leaky, allowing the influx of antibiotics. This could account for its regulation by MarA, and also by BaeR, again subject to conditions when iron is not limited.

### 5.1.2 Other members of the BaeSR regulon

#### 5.1.2.1 *yicO*, *ygcL*, and *ycaC*

*yicO*, *ygcL*, and *ycaC* have been identified in the literature as being upregulated by BaeR (Baranova and Nikaido 2002, Nishino *et al.* 2005). In this project, we demonstrated the ability of BaeR to bind to the promoters of *yicO and ygcL* (Figure 4.2). Furthermore, we identified putative BaeR recognition sequences in the promoter regions of *yicO*, and *ygcL*, while one with lower conformity had previously been located in the *ycaC* promoter (Table 4.1, Nishino *et al.* 2005). *yicO* appears to be in an operon consisting of *yicONM*. Although nothing is known of their function, based on homology, *yicO* appears to be a xanthine uracil permease. Members of this family are integral membrane transporters whose substrate range has been poorly characterized. *yicN* lacks useful homology, while *yicM* is an uncharacterized MF type transporter. *ygcL* appears to be in an operon containing *ygcLKJIHygbTF*. Unfortunately, all of these genes are of unknown function and lack useful homology, with the exception of *ygbT*, which may be involved in DNA repair. Lastly, *ycaC* is thought to be a hydrolase based on sequence homology and crystal structure (Colovos *et al.* 1998). Its substrate has not been identified

and nothing is known of its function. The inclusion of these genes in the BaeR regulon highlights that this pathway may have functions apart from the envelope stress response, possibly in metabolism. We note that the transporters *yicO* and *yicM* add to the list of integral membrane transporters controlled by BaeR, which we expanded to include *cmtAB, yliABCD,* and *fecABCD.* It is possible that by altering the expression profile of cellular transporters, the BaeR pathway mediates adaptation to envelope stress, possibly by functionally complementing other transporters unable to function in a particular environmental condition.

#### 5.1.2.2 *mdtA, acrD,* and *spy*

The canonical members of the BaeSR regulon all have promoters containing BaeR consensus sequences which are very well conserved, and are highly inducible by BaeR overexpression, *baeS1*::Tn*10cam,* and chemical inducers (Baranova and Nikaido 2002, Raffa and Raivio 2002, Hirakawa *et al.* 2005, Nishino *et al.* 2005). *mdtABC* encode a RND type multidrug efflux pump, while *mdtD* is a additional MF type efflux pump (Nagakubo *et al.* 2002). *acrD* also encodes a RND type multidrug efflux pump (Aires and Nikaido 2005). Together these two multidrug efflux pumps mediate BaeR inducible resistance to the antibiotics novobiocin and several β-lactam derivatives, and bile salts, which are common in the gastrointestinal tract (Hirakawa *et al.* 2003a, Hirakawa *et al.* 2003b). These two regulon members epitomize the characterization of the BaeSR pathway as an envelope stress response. The presence of two distinct multidrug efflux pumps in the BaeSR regulon raises the possibility that this pathway responds to envelope stress resulting from toxic compounds partially by increasing the rate of their efflux.

The function of *spy* is more obscure, with its only known homologue being *cpxP*. CpxP is known to be an inhibitory protein, repressing expression of the Cpx regulon, possibly through an association with the sensor kinase CpxA. Proteolytic cleavage of CpxP occurs upon pathway induction (Raivio *et al.* 2000, Buelow and Raivio, 2005). While it has been demonstrated that Spy does not exert an inhibitory effect on the expression of the Cpx regulon and that a deletion of it does not appear to result in increased sensitivity to indole, it is possible that it functions as a repressor of the BaeSR pathway in a manner analogous to the action of CpxP in the CpxAR pathway (Raivio 2000, Raffa and Raivio 2002).

## 5.2 Regulation by BaeSR

### 5.2.1 The BaeR consensus sequence

Recently, a consensus binding sequence was proposed for BaeR, and this sequence was determined to be required for BaeR dependent regulation of *mdtA* and *acrD* (Nishino *et al.* 2005). We identified sequences with homology to the proposed BaeR consensus sequence in all of the promoter regions to which we observed BaeR binding, *cmtB, yicO,* and *ygcL*, and in one, that of *fecI,* for which we did not observe binding (Figure 4.2, Table 4.1). The region of similarity to the consensus sequence identified in the *fecI* promoter was the weakest identified in this study, containing 61% identity and being one bp shorter than the consensus sequence. It is possible that as a result of this BaeR is unable to bind to it under the *in vitro* conditions our experiments used. Also, it has been noted in the literature that BaeR co-regulates *mdtA* and *acrD* in conjunction

with CpxR (Hirakawa *et al.* 2005). It is possible that BaeR is unable to regulate *fec* in the absence of a co-regulator, such as Fur or MarA, both of which are known to downregulate the *fecABCDE* operon (Hantke 2001, Barbosa *et al* 2000).

We noted a contrast between the canonical BaeR regulon members *mdtA, spy* and *acrD,* which have highly conserved BaeR consensus sequences, and the novel regulon members *cmtB*, *ygcL*, and *yicO*, whose consensus sequences are less well conserved, in that the well characterized regulon members appear to be under BaeR inducible control to a much greater degree. In response to BaeR overexpression, expression of a *mdtA-lacZ* reporter construct was reported to be increased by approximately 45 fold as measured by β-galactosidase assay (Baranova and Nikaido, 2002). In comparison, in the same experiment, the *yicO* and *ygcL* promoter regions were found to confer 4-5 and 4.5 fold increases in expression of *lacZ* in response to BaeR over-expression, respectively. When we looked at the induction of our λ*plac*Mu53 insertion located in *mdtA* we observed 2.6 fold induction by indole, 6.9 fold by myricetin, and 11.6 fold by $NaWO_4$. This observation is also noted in the literature, wherein indole has been reported to increase the expression of *acrD* 5 fold, that of *mdtA* 3.9 fold, and that of *spy* nearly 7.5 fold as measured by β-galactosidase assay (Hirakawa *et al.* 2005, Raffa and Raivio 2002). For novel regulon member *cmtB*, we reported that we were unable to observe induction of our λ*plac*Mu53 insertion (located in *yggC*) in response to the BaeSR inducing signals indole, myricetin and $NaWO_4$. The putative regulon member *ycaC* has also been noted to contain a very divergent consensus sequence, with only 50% identity (Nishino *et al.* 2005). Using RT-PCR, fold changes of 640, 490 and 35 were observed for *spy, mdtA,* and *acrD* following BaeSR over-expression. *ycaC* however, was observed to have only 7.3 fold

induction in the same experiments (Nishino *et al.* 2005). Therefore, our data and the literature suggest that the strength of BaeR regulation depends in part on the conformity of the promoter to the BaeR consensus sequence, with the canonical regulon members, which have highly conserved consensus sequences, being highly inducible and the novel regulon members, which have less well conserved consensus sequences, being less inducible under the conditions studied.

It is possible that the degree of homology to the consensus sequence determines the extent to which BaeR recognizes promoters, and in turn the extent to which the genes are BaeSR inducible. This idea is supported by our EMSA studies, in which we observed BaeR binding to the *acrD* promoter at lower protein concentrations than were required for the promoters of *yicO*, *ygcL* and *cmtB*, 0.8μM compared to 1.5μM respectively (figure 4.2). This observation can be extended to the *mdtA* promoter as well, which has been reported to be bound by BaeR at protein concentrations of 0.5μM (Hirakawa *et al.* 2005).

It should also be noted that while the BaeR binding sites experimentally determined by DNaseI protection assays of the *acrD* and *mdtA* promoters have shown a degree of overlap with the consensus sequence determined by sequence alignment of promoter regions, this was not absolute, and in the case of *mdtA* was for only 5bp of the 18bp consensus sequence. Conservation of the consensus sequence is above 94% for *mdtA*, *acrD*, and *spy*, but DNaseI protection assays should be carried out on *spy*, *cmtB*, *yicO* and *ygcL* to further characterize the utility of this sequence for predicting regulation by BaeR.

### 5.2.2 Mechanism of BaeR regulation

It has been reported that *mdtABCD* and *acrD* are regulated by the binding of BaeR and CpxR to the same promoter, resulting in BaeR playing a major role in the activation of expression from these promoters, and CpxR modulating the effect of BaeR (Hirakawa *et al.* 2005). This provides a model of BaeR regulation that we suggest also occurs in conjunction with unknown regulators at the promoters of novel regulon members *cmtBAyggPDFC, ybiKyliABCD* (perhaps at the promoter of an unidentified regulator) and *fecABCD* (perhaps at the *fecIR* promoter). This possibility may account for the lack of induction we observed with chemical inducers of the BaeSR pathway, and the lack of BaeR~P binding to the *fecIR* and *ybiK* promoters despite the presence of a putative BaeR consensus sequence in the *fecIR* promoter. It is possible that this requirement for a second regulator can be overcome by particularly strong activating conditions, such as BaeR overexpression, or the *baeS1*::Tn*10cam* allele.

### 5.2.3 BaeSR activating signals

While the environmental signal that activates the BaeSR pathway is unknown, several inducers are used in the literature, which include spheroplasting, PapG overexpression, and most commonly, indole. The activating signals we utilized in this thesis can be divided into two groups: a genetic background containing the BaeS gain of function mutation *baeS1*::Tn*10cam,* and the chemical inducers, which include indole, myricetin and $NaWO_4$.

#### 5.2.3.1 *baeS1*::Tn*10cam*

After identifying λp*lac*Mu53 insertions in genes that responded to the overproduction of BaeR, we examined the effect of the *baeS1*::Tn*10cam* allele on the insertions in an otherwise wildtype background. We found that for each strain this resulted in a phenotype similar to that observed in the presence of BaeR overexpression (Figure 3.7). For the λp*lac*Mu53 insertions in *mdtA, yggC,* and *yliA* we observed fold inductions of 2.3, 7.7, and 60.8 fold respectively, while the insertions in *fecD* and near *yjgX* showed folds inhibition of 0.6 and 0.4 respectively. We must consider why the large amount of induction we observed for *yggC* and *yliA* was not reported by Nishino *et al.* in their microarray study of gene induction in response to BaeR overexpression (2005). Several possibilities can be considered. It should be noted firstly that the strain background used in the microarray study was W3110, whereas we utilized MC4100. Also, they used a cutoff of 4.0 fold induction to be considered a significant increase, and 0.25 fold to be considered a significant decrease. This second point explains why *fecD* and *yjgX* were not detected in their experiment. It is possible that the mode of action by *baeS1*::Tn*10cam,* which has not been defined, is significant, and explains why we observe such high induction of *yggC* and *yliA* while it was not observed to be as high in BaeR overexpression studies.

The *baeS1*::Tn*10cam* allele was created by the insertion of a Tn*10cam* cassette into the fifth codon of *baeS* (Raffa and Raivio 2002). Two possibilities have been suggested for its effect, firstly that it produces a constitutively active form of BaeS, as well as possibly overexpressing BaeR. I note that the first would entail the second as the insertion would be expected to decouple *baeSR* from its wild type regulation as part of

the *mdtABCDbaeSR* operon. Also, there is a possibility of an internal autoregulated promoter, as *baeS* was found to be induced to a greater extent than the upstream *mdtD* as a result of BaeR overexpression, the opposite of what we would expect from the location of *baeSR* at the 3' end of a larger operon (Nishino *et al.* 2005). Also if BaeS were to be stuck in a conformation where it were active as a kinase, it would be unable to function in its capacity as a phosphatase to dephosphorylate BaeR. This in particular has the potential to result in greater induction of the pathway than BaeR overexpression, as the entire increased pool of cellular BaeR would be phosphorylated, while overexpression of a response regulator is thought to activate its pathway as a result of a small fraction of the total being phoshorylated in the absence of pathway activation (Aguilar *et al.* 2001). It has also been reported that the dephosphorylation of BaeR~P is rapid, possibly due either to high phosphatase activity of BaeS or possibly intrinsic instability of BaeR~P (Yamamoto *et al.* 2005). This observation lends further credence to the suggestion that a constitutively active BaeS may be a stronger inducing signal than BaeR overexpression.

It is possible that promoters with sequences closely matching the consensus sequence such as those of *acrD, mdtA,* and *spy* have a greater affinity for BaeR~P, out competing those of promoters with less conformative consensus sequences, such as *cmtBAyggPFDC* or a putative BaeR regulated regulator of *ybiKyliABCD*. When greater amounts of BaeR~P are available in a *baeS1*::Tn*10cam* strain, these promoters are activated to a greater extent, possibly even overcoming the need for additional activators.

### 5.2.3.2 Chemical activating signals

To date, the most widely used activating signal for the BaeSR pathway has been the chemical indole. It has been reported that treatment of cells with 2mM indole increases the expression of the multidrug efflux pumps MdtABCD and AcrD, resulting in an increase in antibiotic resistance (Hirakawa *et al.* 2005). Indole has also been found to increase the expression of the stress protein Spy (Raffa and Raivio 2002). For all three of these genes, this inducibility was found to depend primarily upon the BaeSR system. Indole is a planar heterocyclic molecule that is thought to mediate its toxicity by dissolving into the lipid bilayer, resulting in membrane derangement (Garbe *et al.* 2000). This effect is thought to result in disruption of the semiquinones leading to the production of superoxide radicals and resultant oxidant toxicity. Interestingly, indole resistance has been found to be conferred by the constitutive expression of *ahpC,* a hydroperoxidase which is a member of the OxyR regulon (Garbe *et al.* 2000). OxyR activates the expression of a number of genes involved in the response to oxidative stress, including *gor* (Toledano *et al.* 1994). *gor* encodes the protein glutathione oxidoreductase, which maintains high levels of glutathione, widely thought to protect cells from oxidative damage, in the cytoplasm (Becker-Hapak and Eisenstark 1995). In our screen, we identified the *yliABCD* operon as being upregulated by BaeR. Since *yliABCD* is known to encode a glutathione uptake system, this result may suggest a role for BaeSR in response to oxidative stress (Suzuki *et al.* 2005).

Another possibility is raised by the observation that the efflux of indole from *E. coli* is mediated by the AcrEF multidrug efflux pump (Kawamura-Sato *et al.* 1999). Several studies have indicated that *acrEF* is not regulated by BaeR (Hirakawa *et al.* 2003,

Nishino *et al.* 2005). This combined with our observation that the novel inducers myricetin and $NaWO_4$ induce BaeR dependent expression of *mdtA* to a much greater extent, may suggest that indole activation of BaeSR may not be a primary effect but rather dependent on a similarity between the oxidative stress effects indole would exert upon the envelope and the actual BaeSR activating signal (Figure 3.10). However it has been strongly suggested that other efflux pathways for indole exist, and the growth of *baeR* mutants was found to be affected by indole to a greater extent than a wildtype strain, leaving this an open question (Raffa and Raivio 2002).

Indole, despite its toxic effects, has also been reported to be involved in extracellular signaling (Wang *et al.* 2001). Indole is produced by *E. coli* as a product of tryptophan catabolism by the *tnaA* gene, and accumulates in stationary phase cultures (Snell 1975). This presence of indole in the medium was found to increase the expression of *gabT, astD,* and *tnaB*, genes involved in the degradation and transport of amino acids (Wang *et al.* 2001). This hints that the regulon of BaeSR may include members that are not clearly involved in the envelope stress response but rather with metabolic functions. This prediction was confirmed by our identification of the *cmtBAyggPFDC* operon which appears to include a PTS type transport system, an oxidoredustase, and a kinase.

The novel inducing cues myricetin and $NaWO_4$ were first implicated as associated with the BaeSR pathway by phenotypic microarray analysis. It was demonstrated that a strain containing a deletion of *baeSR* had increased sensitivity to myricetin, gallic acid, nickel chloride, and sodium tungstate (Zhou 2003). Work completed in our lab showed that of these myricetin, gallic acid, and $NaWO_4$ were able to induce the BaeSR pathway and in this thesis, I demonstrate BaeR dependent induction of *mdtA* by both myricetin and

$NaWO_4$ (Figure 3.10). Myricetin is a plant flavonoid which has been demonstrated to have antimicrobial activity against a variety of antibiotic resistant bacteria, including medically relevant multidrug resistant *Burkholderia cepacia* and vancomycin resistant enterococci (Xu and Lee, 2001). This antimicrobial activity was thought to be due to an effect on protein synthesis as *B. cepacia* incubated with 256μg/mL myricetin was unable in incorporate $^3$H-serine, while $^3$H-thymidine incorporation continued to occur at this concentration. It is puzzling why a compound which affects protein synthesis would induce the BaeSR envelope stress response, however we note that myricetin is a planar multicyclic compound, and it is possible it could insert into the membrane, and cause disruption of the envelope as well.

The toxic effect of $NaWO_4$ appears to be due to the ability of tungstate to act as an analogue of molybdate (Rech *et al.* 1996). Tungstate is able to bind with high affinity to ModA, the periplasmic binding protein responsible for molybdate uptake. Since molybdenum is an essential trace metal required as a cofactor for several enzymes involved in anaerobic respiration and energy generation, blocking its uptake has a toxic effect. It is tempting to speculate that tungstate binding to ModA may lead to increased expression of *modA* as the cell attempts to increase its uptake of molybdate. This could lead to the formation of protein aggregates in the periplasm similar to the overexpression of *papG* which is also known to induce the pathway. Alternatively, a disruption of the cellular processes requiring molybdenum may serve as the BaeSR activating signal.

We were unable to observe any induction of the λp*lac*Mu53 insertions located in *yggC, yliA, fecD,* and near *yjgX* by the chemical inducers we used in this study. It has been noted in the literature that BaeR and CpxR co-regulate *mdtABCD* and *acrD* by

binding to the same promoter (Hirakawa *et al.* 2005). We suggest a similar mechanism is in operation at the promoters of the genes we have identified, or in the case of *yliA, fecD* and *yjgX*, whose promoters we did not observe binding to, possibly at the promoters of another regulator of them. It is possible that these regulators may require additional activating cues before becoming active, explaining our observation of no induction via chemical inducers of the BaeSR pathway. While we have not investigated this possibility, potential candidates may include CysB, known to induce *ybiKyliABCD,* and Fur, a known repressor of both *fecABCDE* and *fecIR.* Strong activating signals, such as BaeR overexpression, or the *baeS1*::Tn*10cam* allele, may overcome this requirement. If BaeSR is involved in regulating the expression of genes involved in cellular metabolism, or in remodeling the envelope in response to stressors, a mechanism such as this may allow for fine-tuning the expression of regulon members, restricting it to times when it is needed both due to stressful conditions, and when it is metabolically appropriate.

**5.3 Future Directions**

The study described in this thesis opens up a variety of directions for novel and continued research. We identified a number of genes, *yggC, yliA, fecD* and *yjgX,* as being BaeSR regulated based on their response to BaeR overexpression and the *baeS1*::Tn*10cam* allele. However we were unable to observe any induction of these genes by inducers of the BaeSR pathway such as indole, myricetin, and $NaWO_4$. In order to further clarify the regulation of these genes by BaeR, I suggest the construction of *lacZ* reporter strains wherein a *lacZ* reporter gene is fused to the promoter of the operon in

question located in the λ recombination site λ*att*. These strains would have two potential advantages over the λp*lac*Mu53 insertion strains first identified in my study. It is notable that three of the insertions identified are located in genes that appear to be members of operons. In all of these cases, the insertion is located in a gene towards the end of the operon, with *yggC* being the sixth gene, *yliA* the second, and *fecD* being the fourth. The last insertion we located is in an intergenic region facing the *yjgX* gene. It is known that as one moves towards the distal ends of an operon the genes are expressed to a lesser degree. Therefore it is possible that if we examined the expression of the first operon member we would be able to observe induction by BaeSR inducing cues. These strains would also allow us to examine induction in an otherwise wildtype background. The λp*lac*Mu53 insertions we isolated would necessarily be null mutations of the gene they are located in, and quite possibly of any operon members further downstream as well due to polar effects. The construction of *lacZ* reporters located in the locus λ*att* would allow us to look at induction of these genes in the absence of any secondary effects arising from null mutations in them. As there are indications in the literature that the expression of *ybiKyliABCD* is greater in minimal media and sulphur limitation, and as it is known that the expression of efflux pumps is affected by media composition, the induction of these strains should be examined in M63 minimal medium as well as LB (Parry and Clark 2002, Bailey *et al.* 2006). Also since regulon members *mdtABCD* and *acrD* confer resistance to bile salts, which would be common in the anaerobic environment of the intestine, it would be interesting to examine the induction of regulon members under anaerobic conditions (Hirakawa *et al.* 2003).

The characterization of myricetin and $NaWO_4$ as strong inducers of the BaeSR pathway gives us a powerful tool to identify further members of this system. I would suggest using microarray technology to compare the expression profiles of MC4100 cells in the presence and absence of $NaWO_4$ since as a powerful activating signal the use of $NaWO_4$ may allow us to identify BaeSR regulated genes that may be missed using a weaker signal such as indole, while at the same time avoiding utilizing unnatural signals such as BaeR over expression. This experiment would be expected to identify genes regulated in response to $NaWO_4$, but would not indicate whether this effect was mediated by BaeR. To clarify this, these results should be compared to a second set of microarray data showing the expression profile of MC4100 Δ*baeR* in response to $NaWO_4$. We would expect genes which are in fact regulated by BaeR in response to $NaWO_4$ to display no difference in expression in this second set of data. A second method to expand our knowledge of the BaeSR regulon would be to use a microarray to compare the expression of genes in TR776 (MC4100 *baeS1::*Tn*10cam*) with MC4100. As *baeS1::*Tn*10cam* strongly induces the BaeSR pathway, this should lead to the identification of novel regulon members.

As the BaeSR pathway represents a third envelope stress response pathway, we expect that many of its members will be involved in the cellular reaction to stressful conditions. In order to test this we could compare growth curves for MC4100 and the λ*plac*Mu53 insertions created in this study under a variety of stressful and non-stressful conditions. These should include the substrates of the MdtABCD, novobiocin and deoxycholate, and the BaeSR inducing cues indole, myricetin and $NaWO_4$. It is possible that growth curves would yield more information than standard MIC studies if the growth

of these mutants would be affected while the overall MIC to the compound in question would remain unchanged.

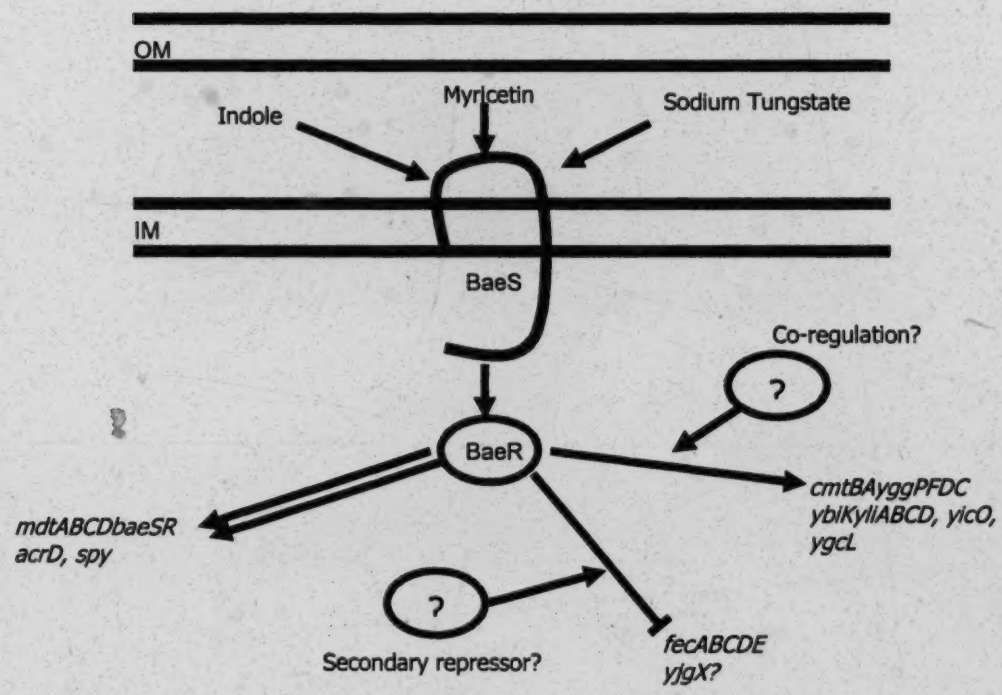
Our identification of a λp*lac*Mu53 insertion in *mdtA* also gives us a useful tool to examine the role of the Spy protein, which is currently unknown. Spy, the most highly inducible member of the BaeSR regulon, has homology only to CpxP, the inhibitor, and most highly inducible member of, the CpxAR pathway (Nishino *et al.* 2005, Raivio *et al.* 2000). Previous research has demonstrated that *spy* overexpression did not inhibit expression of *cpxP* and *degP,* members of the CpxAR regulon (Raivio *et al.* 2000). Since these genes are not regulated by BaeSR, it remains possible that Spy plays a role in the BaeSR pathway similar to that of CpxP in the CpxAR pathway (Raffa and Raivio 2002). To test this possibility, I suggest transforming the strain containing the λp*lac*Mu53 insertion in *mdtA* (CO4) with the Spy overexpression plasmid, ptrc99A:*spy*. The inducibility by $NaWO_4$ of these two strains should then be examined, in the presence and absence of IPTG induction of Spy overexpression. If Spy does act as an inhibitor of the BaeSR pathway, we would expect to be unable to induce *mdtA* when it is overexpressed.

Finally, our identification of putative BaeR consensus sequences in the promoters of *cmtB, yicO,* and *ygcL* should be confirmed by use of DNaseI protection assays. This would also have the benefit of further clarifying the sequence and utility of the proposed BaeR consensus sequence.

In conclusion, this project has been successful in providing evidence that several novel genes, *yggC, yliA, fecD,* and *yjgX* are BaeR regulated, resulting in increased understanding of this pathway (Figure 5.1). This study provides several new avenues of

**Figure 5.1 A revised view of the BaeSR pathway**

This thesis presented evidence for the upregulation of *yggC*, and *yliA,* and the downregulation of *fecD* and *yjgX* by BaeSR. As we were unable to observe induction of these genes by chemical inducing cues of the BaeSR pathway, it is possible that regulation of them by BaeR depends partially upon secondary, unidentified regulatory proteins. We have also identified myricetin and NaWO4 as strong inducing cues of BaeSR, providing increased knowledge about the activation of this pathway.


OM
Indole
Myricetin
Sodium Tungstate
IM
BaeS
Co-regulation?
?
BaeR
mdtABCDbaeSR
acrD, spy
cmtBAyggPFDC
ybiKyliABCD, yicO,
ygcL
?
Secondary repressor?
fecABCDE
yjgX?

research in order to confirm and expand upon these results, as detailed in section 5.3. We have also provided evidence demonstrating specific BaeR binding to the promoter regions of *cmtB*, *yicO*, and *ygcL*, confirming their identity as BaeR regulated genes, and have experimentally confirmed the BaeR binding site in the *acrD* promoter reported by Hirakawa *et al.* (2005). We have characterized the induction of *mdtA* by the novel inducing cues myricetin and $NaWO_4$, showing them to be stronger inducers of the pathway than indole, and therefore useful tools for screening for additional BaeSR regulated genes. The inclusion of a number of transporters and membrane associated proteins in the BaeSR regulon suggests that the BaeSR pathway may play a role in cellular metabolism in addition to its role in the envelope stress response, or that it may be involved in remodeling the expression profile of membrane localized proteins in response to stressors. Further study along the lines indicated is sure to expand our knowledge of the physiological role of BaeSR even further.

## 6. References

**Ades, S.E., Connolly, L.E., Alba, B.M., and Gross, C.A.** 1999. The *Escherichia coli* $\sigma^E$-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-$\sigma$ factor. Genes Dev. **13**:2449-2461

**Aguilar, P.S., Hernandez-Arriaga, A.M., Cybulski, L.E., Erazo, A.C., and Mendoza, D.** 2001. Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilus*. EMBO **20(7):** 1681-1691

**Ahmed, M., Borsch, C.M., Taylor, S.S., Vázquez-Laslop, N. and Neyfakh, A.A.** 1994. A protein that activates expression of a multidrug efflux transporter upon binding the transporter substrates. J Biol Chem. **269(45)**:28506-28513

**Aires, J.R., and Nikaido, N.** 2005. Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD Multidrug efflux transporter of *Escherichia coli*. J Bacteriol. **187(6)**:1923-1929

**Alba, B.M., Leeds, J.A., Onufryk, C., Lu, C.Z., and Gross, C.A.** 2002. DegS and YaeL participate sequentially in the cleavage of RseA to activate the $\sigma^E$ –dependent extracytoplasmic stress response. **16(16):** 2156-2168

**Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403-410

**Andersen, C., Koronakis, E., Bokma, E., Eswaran, J., Humphreys, D., Hughes, C., Koronakis, V.** 2002. Transition to the open state of the TolC periplasmic tunnel entrance. Proc. Nat. Acad.Science, **99(17)**:11103-11108

**Appleby, J.L., Parkinson, J.S., and Bourret, R.B.** 1996. Signal transduction via the multi-step phosphorelay: Not necessarily a road less traveled. Cell. **86:** 845-848

**Avila-Sakar, A.J., Misaghi, S., Wilson-Kubalek, E.M., Downing, K.H., Zygurkaya, H., Nikaido, H., and Nogales, E.** 2001. Lipid-Layer crystallization and preliminary three-dimensional structural analysis of AcrA, the periplasmic component of a bacterial multidrug efflux pump. J Structural Biol. **136:** 81-88

**Bailey, A.M., Webber, M.A., and Piddock, L.J.V.** 2006. Medium plays a role in determining expression of *acrB, marA,* and *soxS* in *Escherichia coli.* Antimicrobial Agents and Chemotherapy. **50(3):** 1071-1074.

**Baranova, N. and Nikaido, H.** 2002. The BaeSR two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. J Bacteriol. **184(15):** 4168-4176

**Barbosa, T.M., and Levy, S.B.** 2000. Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. J Bacteriol. **182(12):** 3467-3474

**Batchelor, E., Walthers, D., Kenney, L.J., and Goulain, M.** 2005. The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins OmpF and OmpC. J Bacteriol. **187(16):** 5723-5731

**Becker-Hapak, M., and Eisenstark, A. 1995.** Role of *rpoS* in the regulation of glutathione oxidoreductase (gor) in *Escherichia coli.* FEMS Microbiology letters. **134:** 39-44

**Bilwes, A.M., Alex. L.A., Crane, B.R. and Simon, M.I.** 1999. Structure of CheA, a Signal-transducing Histidine Kinase. Cell. **96:** 131-141

**Birck, C., Malfois, M., Svergun, D., and Samama JP.** 2002. Insights into signal transduction revealed by the low resolution structure of the FixJ response regulator. J. Mol. Biol. **321:** 447-457

**Birck, C., Mourey, L., Gouet, P. Fabry, B., Schumacher, J., and Rousseau, P.** 1999. Conformational changes induced by phosphorylation of the FixJ reciever domain. Structure. **7:** 1505-1515

**Bjarnason J., Southward C.M., and Surette, M.G.** 2003. Genomic profiling of iron-responsive genes in *Samonella enterica* serova Typhimurium by high-throughput screening of a random promoter library. J Bacteriol. **185:** 4973-4982

**Borek, D., Michalska, K., Brzezinski, K., Kisiel, A., Podkowinski, J., Bonthron, D.T., Krowarsch, D., Otewski, J., and Jaskolski, M.** 2004. Expression, purification and catalytic activity of *Lupinus luteus* asparagines β-amidohydrolayse and its *Escherichia coli* homolog. Eur. J. Biochem. **271:** 3215-3226

**Brass, J.M., Higgins, C.F., Foley, M., Rugman, P.A., Birmingham, J., and Garland, P.B.** 1986. Lateral Diffusion of proteins in the periplasm of *Escherichia coli.* J Bacteriol. **165:** 787-794

**Bremer, E., Silhavy, T.J. and Weinstocks, G.M.** 1985 Transposable λ p*lac*Mu bacteriophage for creating *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli.* J Bacteriol. **162(3):** 1092-1099

**Buchanan, S.K.,** 2005. Bacterial metal detectors. Mol. Microbiol. **58(5):** 1205- 1209

**Buelow, D.R., and Raivio, T.L.** 2005. Cpx signal transduction is influenced by a conserved N-Terminal domain in the novel inhibitor CpxP and the Periplasmic Protease DegP. J Bacteriol. **187(19):**6622-6630

**Casadaban, M.J.** 1976 Transposition and fusion of *lac* genes to selected promoters in *Escherichia coli* using bacteriophages lambda and Mu. J Mol Biol. **104:** 541-555

**Caetano-Anolles, G.** 1993. Amplifying DNA with arbitrary oligonucleotide primers. *PCR Methods Appl* **3**: 85–94

**Cohen, S.P., Levy, S.B., Foulds, J., and Rosner, J.L.** 1993. Salicylate induction of antibiotic resistance in *Escherichia coli*: activation of the *mar* operon and a *mar*-independent pathway. J Bacteriol. **175:** 7856-7862

**Connolly, L., Peñas, A.D.L., Alba, B.M., Gross, C.A.** 1997. The response to extracytoplasmic stress in *Escherchia coli* is controlled by partially overlapping pathways. Genes Dev. **11:** 2012-2021

**Collinet B., Yuzawa, H., Chen, T., Herrera, C., and Missiakas, D.** 2000. RseB binding to the periplasmic domain of RseA modulates the RseA:$\sigma^E$ Interaction in the cytoplasm and the availability of $\sigma^E$ RNA polymerase. J Biol Chem. **275(43):**33898-33904

**Colovos, C., Cascio, D., and Yeats, T.O.** 1998. The 1.8A crystal structure of the ycaC gene product from *Escherichia coli* reveals an octameric hydrolase of unknown specificity. Structure. **6(10):** 1329-1337

**Cosma, C.L., Danese, P.N., Carlson, J.H. Silhavy, T.J. and Snyder, W.B.** 1995. Mutational activation of the Cpx signal transduction pathway of *Escherichia coli* suppresses the toxicity conferred by certain envelope-associated stresses. Mol Microbiol. **18(3):** 491-505

**Danese, P.N. and Silhavy, T.J.** 1997. The $\sigma^E$ and Cpx signal transduction systems control the synthesis of periplasmic protein-folding enzymes in *Escherchia coli.* Genes Dev. **11:** 1183-1193

**Danese, P.N. and Silhavy, T.J.** 1998. CpxP, a stress-combative member of the Cpx regulon. J Bacteriol. **180(4):** 831-839

**Danese, P.N., Snyder, W.B., Cosma, C.L., Davis, L.J.B. and Silhavy, T.J.** 1995. The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. Genes Dev. **9:** 387-398

**Dartigalongue, C., Missiakas, D., and Raina, S.** 2001. Characterization of the *Escherichia coli* :$\sigma^E$ Regulon. J Biol Chem. **276(24):**20866-20875

**De Las Peñas, A., Connolly, L., Gross, C.A.** 1997. The $\sigma^E$–mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of $\sigma^E$. Mol Microbiol. **24(2):** 373-385

**De Wulf, P., McGuire, A.M., Liu, X., and Lin, E.C.C.** 2002. Genome-wide profiling of promotor recognition by the two-component response regulator CpxR-P in *Escherichia coli.* J Biol Chem. **277(29):** 26652-26661

**DiGiuseppe, P.A. and Silhavy, T.J.** 2003. Signal detection and target gene induction by the CpxRA two-component system. J Bacteriol. **185(8):** 2432-2440

**Ding, H., Hidago, E., Demple, B.** 1996. The redox state of the [2Fe-2S] clusters in SoxR protein regulates its activity as a transcription factor. J Biol Chem **271:** 33173-33175

**Dinh, T., Paulsen, I.T. and Saier, M.H.** 1994. A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of Gram-negative bacteria. J Bacteriol. **176(13):** 3825-3831

**Dong, J., Icchi, S., Kwan, H.S., Lu, Z. and Lin, E.C.C.** 1993. The deduced amino-acid sequence of the cloned *cpxR* gene suggests the protein is the cognate regulator for the membrane sensor, CpxA, in a two-component signal trnsduction system of *Escherichia coli.* Gene. **136:** 227-230

**Elkins, CA, and Nikaido, H.** 2003. Chimeric analysis of AcrA function reveals the importance of its C-terminal domain in its interaction with the AcrB multidrug efflux pump. J Bacteriol. 2003 **185(18):** 5349-56

**Erickson, J.W. and Gross, C.A.** 1989. Identification of the $\sigma^E$ subunit of *Escherichia coli* RNA polymerase: a second alternative $\sigma$ factor involved in high-temperature gene expression. Genes Dev. **3:** 1462-1471

**Fahey, R.C., Brown, W.C., Adams, W.B., and Worsham, M.B.** 1975. Occurrence of glutathione in bacteria. J. Bacteriol. **133:** 1126-1129.

**Fralick, J.A.** 1996. Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli.* J Bacteriol. **178(19):** 5803-5805

**Garbe, T.R., Kobayashi, M., and Yukawa, H.** 2000. Indole-inducible proteins in bacteria suggest membrane and oxidant toxicity. Arch Microbiol. **173:** 78-82

**Garzón, A. and Parkinson, J.S.** 1996. Chemotactic signaling by the P1 phosphorylation domain liberated from the CheA histidine kinase of *Escherichia coli.* J Bacteriol. **178(23):** 6752-6758

**Goudreau, P.N. and Stock, A.M.** 1998. Signal transduction in bacteria: molecular mechanisms of stimulus-response coupling. Curr Opin Microbiol. **1:** 160-169

**Gouet, P., Fabry, , B., Guillet, V., Birck, C., Mourey, L., Kahn, D., and Samama, J.P.** 1999. Structural transitions in the FixJ reciever domain. Structure. **7:** 1517-1526

**Greenberg, J.T. and Demple, B.** 1986. Glutathione in *Escherihia coli* is despensible for resistance to $H_2O_2$ and gamma radiation. J Bacteriol. **168(2):** 1026-1029

**Grimsly, J.K., Tjalkens, R.B., Strauch, M.A., Bird, T.H., Spiegelman, G.B., Hostomsky, Z., Whiteley, J.M., and Hoch, J.A.** 1994. Subunit composition and domain structure of the Spo0A sporulationt factor of *Bacillus subtilus*. J Bio Chem. **269(24):** 16977-16982

**Gross, C.A.** 1996. Function and regulation of the heat-shock proteins. P. 1382-1399. In F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaecter, and H. Umbarger (ed.). *Escherichia coli and Salmonella: Cellular and Microbiology* 2nd ed. ASM Press, Washington, D.C.

**Hagenmaier, S., Stierhof, Y. and Henning, U.** 1997. A new periplasmic protein of *Escherichia coli* which is synthesized in spheroplasts but not in intact cells. J Bacteriol. **179(6):** 2073-2776

**Hakenbeck, R. and Stock, J.B.** 1996. Analysis of two-component signal transduction systems involved in transcriptional regualtion. Methods Enzymol. **273:** 281-300

**Hantke, K.** 2001. Iron and metal regulation in bacteria. Curr Opin Microbiol. **4:** 172-177

**Harrod, A.C., Yang, X., Junker, M., Reitzer, L.** 2004. Evidence for a second interaction between the regulatory amino-terminal and central output domains of the response regulator NtrC (Nitrogen Regulator I) in *Escherichia coli.* J. Biochem. **279(4):** 2350-2359

**Harshey, R.M. and Toguchi A.** 1996. Spinning tails: Homologies among bacterial flagellar systems. Trends in Microbiology. **4(6):** 226-231

**Helling, R.B., Janes, B.K., Kimball, H., Tran, T., Bundesmann, M., and Check, P.** 2002. Toxic waste disposal in *Escherichia coli.* J Bacteriol. **184:** 3699-3703.

**Hershberg, R., Altuvia, S., and Margalit, H.** 2003. A survey of small RNA-encoding genes in *Escherichia coli.* Nucleic Acids Research. **31(7):** 1813-1820

**Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., and Yamaguchi, A.** 2005. Indole induces the expression of multidrug exporter genes in *Escherichia coli.* Mol Microb. **55(4):** 1113-1126

**Hirakawa, H., Nishino, K., Yamada, J., Hirata, T. and Yamaguchi, A.** 2003a. β-Lactam resistance modulated by the overexpression of response regulators of two-component signal transduction systems in *Escherichia coli.* J Antimicrob Chemother. **52:** 576-582

**Hirakawa, H., Nishino, K., Hirata, T. and Yamaguchi, A.** 2003b. Comprehensive studies of drug resistance mediated by overexpression of response regulators of two-component signal transduction systems in *Escherichia coli.* J Bacteriol. **185(6):** 1851-1856

**Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., and Yamaguchi, A.** 2005. Indole induces the expression of multidrug exporter genes in *Escherichia coli.* Mol Microb. **55(44):**1113-1126

**Hirschman, A., Boukhvalova, M., VanBruggen, R., Wolfe, A.J., and Steward, R.C.** 2001. Active site mutations in CheA, the signal transducing protein kinase of the chemotaxis system in *Escherichia coli.* Biochemistry. **40:** 13876-13887

**Hoch, J., and Silhavy, T.J.** (ed.) 1995. *Two-compomnent signal transduction.* American Society of Microbiology, Washington, D.C.

**Hung, D.L., Raivio, T.L., Jones, C.H., Silhavy, T.J., and Hultgren, S.J.** 2001. Cpx signaling pathway monitors biogenisis and affects assembly and expression of P pili. EMBO J. **20(7):** 1508-1518

**Ireton, K., Rudner, D.Z., Siranosian, K.J., and Grossman, A.D.** 1993. Integration of multiple developmental signals in *Bacillus subtilus* through the Spo0A transcription factor. Genes Dev. **7:** 283-294

**Jones, C.H., Danese, P.N., Pinkner, J.S., Silhavy, T.J., and Hultgren, S.J.** 1997. The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. EMBO J. **16:** 6394-6406

**Kadner, R.J.** 1996. Cytoplasmic Membrane. P. 58-87. In F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaecter, and H. Umbarger (ed.). *Escherichia coli and Salmonella: Cellular and Microbiology* 2nd ed. ASM Press, Washington, D.C.

**Kallenius, G., Mollby, R., Svenson, S.B., Helin, I., Hultberg, H., Cedergren, B., and Winberg, J.** 1981. Occurance of P-fimbriated *Escherichia coli* in urinary tract infections. Lancet. **2(8260-61):**1369-72

**Killmann, H., Benz, R., and Braun, V.** 1996. Properties of the FhuA channel in the *Escherichia coli* outer membrane after deletion of FhuA portions within and outside the predicted gating loop. J. Bacteriol. **178(23):** 6923-6920

**Kim, D., and Forst, S.** 2001. Genomic analysis of the histidine kinase family in bacteria and archea. Microbiology. **147(5):** 1197-1212.

**Kolmar, H., Waller, P.R.H. and Sauer, R.T.** 1996. The DegP and DegQ periplasmic endoproteases of *Escherichia coli:* Specificity for cleavage sites and substrate conformation. J Bacteriol. **178(20):** 5925-5929

**Koronakis, V.** 2003. TolC – the bacterial exit duct for proteins and drugs. FEBS Letters. **555** 66-71.

**Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C.** 2000. Crystal structure of the bacterial membrane portein TolC central to multidrug efflux and protein export. Nature. **405(6784):**914-919

**Kredich, N.M.** 1996. Biosynthesis of Cystine. P. 514-527. In F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaecter, and H. Umbarger (ed.). *Escherichia coli and Salmonella: Cellular and Microbiology* 2nd ed. ASM Press, Washington, D.C.

**Lee, S.Y., C. H.S., Pelton, J.G. Yan, D., Henderson, R.K., King, D.S., Huang, L.S., Kustu, S., Berry, E.A. and Wemmer, D.E.** 2001. Cryatal structure of an activated response regulator bound to its target. Nat. Structur. Biol. **8(1):** 52-56

**Levy, S.B.** 1992. Active efflux mechanisms for antimicrobial resistance. Antimicrob Agents Chemother. **36(4):** 695-703

**Levy, S.B.** 2002. Active efflux, a common mechanism for biocide and antibiotic resistance. Journal of applied Microbiology Symposium Supplement. **92** 65S-71S

**Lewis, R.J., Brannigan, J.A. Muchová, Barák, I. and Wilkinson, A.J.** 1999. Phosphorylated aspartate in the structure of a response regulator protein. J Mol Biol. **294:** 9-15

**Lipinska, B., Sharma, S. and Georgopoulos.** 1988. Sequence analysis and regulation of the *htrA* gene of *Escherichia*: a $\sigma^{32}$-independent mechanism of heat-inducible transcription. Nuc Acids Res. **16(21):** 10053-10067

**Lipinska, B., Zylicz, M. and Georgopoulos, C.** 1990. The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. J Bacteriol. **172(4):** 1791-1797

**Lomovskaya, O., Lewis, K., and Matin, A.** 1995. EmrR is a negative regulator for the *Escherchia coli* multidrug resistance pump EmrAB. J Bacteriol. **177(9):** 2328-2334

**Lonetto, M.A., Brown, K.L., Rudd, K.E. and Buttner, M.J.** 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existance of a subfamily of eubacterial RNA polymerase σ factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. **91:** 7573-7577

**Lukat, G.S., McCleary, W.R., Stock, A.M. and Stock, J.B.** 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. Proc. Natl. Acad. Sci. **89:** 718-722

**Ma, D., Cook, D.N., Alberti, M., Pon, N.G., Nikaido, H. and Hearst, J.E.** 1995. Genes *acrA* and *acrB* encode stress-induced efflux system of *Escherchia coli.* Mol Microbiol. **16(1):** 45-55

**Ma, D., Alberti, M., Lynch, C. Nikaido, H. and Hearst, J.E.** 1996. The local repressor AcrR plays a modulating role in the regulation of *acrAB* genes of *Escherchia coli* by global stress signals. J Bacteriol. **19(1):** 101-112

**Martínez-Hackert, E. and Stock, A.M.** 1997. Structural Relationships in the OmpR Family of Winged-Helix Transcription Factors. J. Mol. Biol. **269:** 301-312

**Mayover, T.L. Halkides, C.J. and Stewart, R.C.** 1999. Kinetic characterization of the CheY phosphorylation reactions: comparision of P-CheA and small-molecule phosphodonors. Biochemistry. **38:** 2259-2271

**McCleary, W.R., Stock, J.B. and Ninfa, A.J.** 1993. Is acetyl phosphate a global signal in *Escherichia coli* ? J Bacteriol. **175(10):** 2793-2798

**McEwen J. and Silverman, P.** 1979. Chromosomal mutations of *Escherichia coli* that alter expression of conjugative plasmid functions. Proc. Natl. Acad. Sci. **77(1):** 513-517

**McEwen, J. and Silverman, P.M.** 1982. Mutations in Genes *cpxA* and *cpxB* alter the protein composition of *Escherichia coli* inner and outer membranes. J Bacteriol. **151(3):** 1553-1559

**McEwen, J., Sambucetti, L., and Silverman, P.M.** 1983. Synthesis of outer membrane proteins in *cpxA cpxB* mutants of *Escherichia coli* K-12. J Bacteriol. **154(1):** 375-382

**Mecsas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J. and Gross, C.A.** 1993. The activity of $\sigma^E$ an *Escherichia coli* heat-inducible σ-factor, is modulated by expression of outer membrane proteins. Genes Dev. **7:** 2618-2628

**Mileykovskaya, E., and Dowhan, W.** 1997. The Cpx two-component signal transduction pathway is activated in *Escherichia coli* mutant strains lacking phosphatidylethanolamine. J. Bacteriol. **179:** 1029-1034

**Missiakas, D. and Raina, S.** 1998. The extracytoplasmic function sigma factors: role and regulation. Mol Microbiol. **28(6):** 1059-1066

**Murakami, S., Nakashima, R., Yamashita, E., and Yamaguchi, A.** 2002. Crystal structure of bacterial multidrug efflux transporter AcrB. Nature. **419:** 587-593

**Nagakubo, S., Nishino, K., Hirata, T. and Yamaguchi, A.** 2002. The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. J Bacteriol. **184(15):** 4161-4167

**Nagasawa, S., Ishige, K., Mizuno, T.** 1993. Novel members of the two-component signal transduction genes in *Escherichia coli.* J Biochem. **114(3):** 350-357

**Nikaido, H.** 1996a. Outer Membrane. P. 29-47. In F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaecter, and H. Umbarger (ed.). *Escherichia coli and Salmonella: Cellular and Microbiology* 2nd ed. ASM Press, Washington, D.C.

**Nikaido, H.** 1996b. Multidrug efflux pumps of Gram-negative bacteria. J Bacteriol. **178(20):** 5833-5859

**Nishino, K. Honda, T., and Yamaguchi, A.** 2005. Genome-wide analyses of *Escherichia coli* gene expression responsive to the BaeSR two-component regulatory system. J Bacteriol. **187(5):** 1763-1772

**Nishino, K. and Yamaguchi, A.** 2001a. Overexpression of the response regulator *evgA* of the two-component signal transduction system modulates multidrug resistance conferred by multidrug resistance transporters. J Bacteriol. **184(4):** 1455-1458

**Nishino, K. and Yamaguchi, A.** 2001b. Analysis of a complete library of putative drug transporter genes in *Escherichia coli.* J Bacteriol. **183(20):** 5803-5812

**Nishino, K. and Yamaguchi, A.** 2002. EvgA of the two-component signal transduction system modulates production of the YhiUV multidrug transporter in *Escherichia coli.* J Bacteriol. **184(8):** 2319-2323

**Nixon, B.T., Ronson, C.W., and Ausubel, F.M.** 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC.* Proc. Natl. Acad. Sci. **83:** 7850-7854

**Ohashi, K., Yamashino, T., and Mizuno, T.** 2004. Molecular basis for promoter selectivity of the transcriptional activator OmpR of *Escerichia coli*: isolation of mutants that can sctivate the non-cognate *kdpABC* promoter. J Biochem. **137:** 51-59

**Okusu, H., Ma, D. and Nikaido, H.** 1996. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherchia coli* multiple-antibiotic-resistance (Mar) mutants. J Bacteriol. **178(1):** 306-308

**Oliver, D.B.** 1996. Periplasm. P. 88-103. In F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaecter, and H. Umbarger (ed.). *Escherichia coli and Salmonella: Cellular and Microbiology* 2nd ed. ASM Press, Washington, D.C.

**Otto, K. and Silhay, T.J.** 2002. Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. Proc. Natl. Acad. Sci. **99(4):** 2287-2292

**Parkinson, J.S.** 1993. Signal transduction schemes of bacteria. Cell. **73:** 857-871

**Paulsen, I.T., Sliwinski, M.K. and Saier, M.H.** 1998. Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificites. J Mol Biol. **277:** 573-592

**Parry, J., and Clark. D.P.** 2002. Identification of a CysB-regulated gene involved in glutathione transport in *Escherichia coli.* FEMS Microbiol Letters. **202:** 81-85

**Penninckx, M.J. and Elskens, M.T.,** 1993. Metabolism and functions of glutathione and thioredoxin pathways in micro-organisms. Adv. Microb. Physiol. **34:** 239-301.

**Powell, B.S., and Kado, C.I.** 1990. Specific binding of VirG to the vir box requires a C-terminal domain and exhibits a minimum concentration threshold. Mol. Microbiol. **4:** 2159-2166

**Prieto-Álamo, M., Jurado, J., Gallardo-Madueño, R., Monje-Casas, F., Holmgren, A., and Pueyo, C.** 2000. Transcriptional regulation of glutaredoxin and thioredoxin pathways and related enzymes in response to oxidative stress. J Biol Chem. **275(18):** 13398-13405.

**Raetz, C.R.H., and Dowhan, W.** 1990. Biosynthesis and function of phospholipids in *Escherchia coli.* J Biol Chem. **265(3):** 1235-1238

**Raffa, R.G. and Raivio, T.L.** 2002. A third envelope stress signal transduction pathway in *Escherichia coli.* Mol Microbiol. **45(6):** 1599-1611

**Raivio, T.L., Laird, M.W., Joly, J.C. and Silhavy, T.J.** 2000. Tethering of CpxP to the inner membrane prevents spheroplast induction of the Cpx envelope stress response. Mol Microbiol. **37(5):** 1186-1197.

**Raivio, T.L. and Silhavy, T.J.** 1999. The $\sigma^E$ and Cpx regulatory pathways: overlapping but distinct envelope stress responses. Curr Opin Microbiol. **2:** 159-165

**Raivio, T.L. and Silhavy, T.J.** 2000. Sensing and responding to envelope stress. In *Bacterial Stress Responses.* ASM Press. 19-32

**Raivio, T.L. and Silhavy, T.J.** 2001a. Regulatory systems: two-component. In *Encyclopedia of Life Sciences*, Nature Publishing Group.

**Raivio, T.L. and Silhavy, T.J.** 2001b. Periplasmic stress and ECF sigma factors. Annu Rev Microbiol. **55:** 591-624

**Rakonjac, J., Milic., M., and Savic, D.J.** 1991. *cysB* and *cysE* mutants of *Escherichia coli* K12 show increased resistance to novobiocin. Mol. Gen. Genet. **228(1-2):** 307-311

**Rech, S., Wolin, C., and Gunsalus, R.P.** 1996. Properties of the periplasmic ModA molybdate-binding protein of *Escherichia coli.* J Biol Chem. **271(5):** 2557-2562

**Rosenberg, E.Y., Bertenthal, D., Nilles, M.L., Bertrand, K.P. and Nikaido, H.** 2003. Bile salts and Fatty acids induce the expression of *Escherichia coli* AcrAB multidrug efflux pump through their interaction with Rob regulatory protein. Mol Microbiol. **48(6):** 1609-1619

**Saito, K., Ito, E., Hosono, K., Nakamura, K. Imai, K., Iizuka, T., Shiro, Y., and Nakamura, H.** 2003. The uncoupling of oxygen sensing, phosphorylation signaling, and transcriptional activation in oxygen sensor FixL and FixJ mutants. Mol Microb. **42(2):** 373-383

**Sauter, A., and Braun, V.** 2004. Defined inactive FecA derivatives mutated in functional domains of the outer membrane transport and signaling protein of *Escherichia coli* K-12. J Bacteriol. **186(16): 5303-5310**

**Shannon, K.P. and French, G.L.** 2004 Increasing resistance to antimicrobial agents of Gram-negative organisms isolated at a London teaching hospital, 1995-2000. J Antimicrob Chemother. **53(5):** 818-825

**Silhavy, T.J., Berman, M.L., and Enquist, L.W.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory

**Slauch, J.M., Russo, F.D., and Silhavy, T.J.** 1991. Suppressor mutations in *rpoA* suggest that OmpR controls transcription by direct interaction with the α subunit of RNA Polymerase. J Bacteriol. **173(23):** 7501-7510

**Snell, E.E.** 1975. Tryptophanase: structure, catalytic activities, and mechanism of action. Adv. Enzymol. Relat. Areas Mol. Biol. **42:** 287-333

**Snyder, W.B., Davis, L.J., Danese, P.N., Cosma, C.L., and Silhavy, T.J.** 1995. Overproduction of NlpE, a new outer membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal transduction pathway. J. Bacteriol. **177:** 4216-4223.

**Sprenger, G.A.** 1993. Two open reading frames adjacent to *Escherichia coli* K-12 transketolase (*tkt*) gene show high similarity to the mannitol phosphotransferase system emzymes from *Escherichia coli* and various Gram-positive bacteria. Biochimica et Biophysica Acta. **1158:** 103-106

**Stock, A., Koshland, D.E., and Stock, J.** 1985. Homologies between the *Salmonella typhimurium* CheY protein and proteins involved in the regulation of chemotaxis, membrane protein synthesis, and sporulation. Proc. Natl. Acad. Sci. **82:** 7989-7993

**Strauch, K.L. and Beckwith, J.** 1988. An *Escherichia coli* mutation prevention degradation of abnormal periplasmic proteins. Proc. Natl. Acad. Sci. **85:** 1576-1580

**Surette, M.G. Levit, M., Liu, Y. Lukat, G. Ninfa, E.G., Ninfa, A. and Stock, J.B.** 1996. Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis. J Biol Chem. **271(2):** 939-945

**Suzuki, H., Koyanagi, T., Izuka, S., Onishi, A., and Kumagai, H.** 2005. The *yli, -B, -C,* and *–D,* genes of *Escherichia coli* K-12 encode a novel glutathione importer with an ATP-Binding cassette. J. Bacteriol. **187(17):** 5861-5867

**Tam, C., Collinet, B., Lau, G., Raina, S., and Missiakas, D.** 2002. Interaction of the conserved region 4.2 of $\sigma^E$ with the RseA Anti-sigma factor. J Biol Chem. **277(30):**27282-27287

**Thanassi D.G., Suh, G.S.B. and Nikaido, H.** 1995. Role of outer membrane barrier in efflux-mediated tetracycline resistance of *Escherichia coli.* J Bacteriol. **177(4):** 998-1007

**Thanassi D.G., Cheng, L.W. and Nikaido, H.** 1997. Active efflux of bile salts by *Escherchia coli.* J Bacteriol. **179(8):** 2512-2518

**Toledano, M.B., Kullik, I., Trinh, F., Baird, P.T., Schneider, T.D., and Storz, G.** 1994. Redox-dependent sift of OxyR-DNA contacts along an extended DNA-binding site: A mechanism for differential promoter selection. Cell. **78:** 897-909

**Tomomori, C., Toshiyuki, T., Dutta, R., Park, H., Saha, S.K., Zhu, Y., Ishima, R., Liu, D., Tong, K.I., Kurokawa, H., Qian, H., Inouye, M., and Ikura, M.** 1999. Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. Nat Struct Biol. **6(8):** 729-734

**Touzé, T., Eswaran, J., Bokma, E., Koronakis, E., Hughes, C., and Koronakis, V.** 2004. Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system. Mol Micorb. **52(2):**697-706

**Van Bambeke, F. Balzi, E. and Tulkens, P.M.** 2000. Antibiotic Efflux Pumps. Biochemical Pharmacology. **60:** 457-470.

**Verhamme D.T., Arents J.C., Postma P.W., Crielaard W., Hellingwerf K.J.** 2002. Investigation of in vivo cross-talk between key two-component systems of *Escherichia coli*. Microbiology. **148(1):** 69-78

**Wang, D., Ding, X. and Rather, P.N.** 2001. Indole can act as an Extracellular Signal in *Escherichia coli.* J Bacteriol. **183(14):** 4210-4216

**Weber, R.F. and Silverman, P.M.** 1988. The Cpx proteins of *Escherichia coli* K12. structure of the CpxA polypeptide as an inner membrane component. J Mol Biol. **203:** 467-478

**Wedel, A., and Kutsu, S.** 1995. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. Genes Dev. **9:**(2042-2052)

**Wei, Y., Lee, J., Smulski, D.R., and LaRossa, R.A.** 2001. Global Impact of *sdiA* Amplification Revealed by Comprehensive Gene Expression Profiling of *Escherichia coli.* J Bacteriol. **183(7):** 2265-227

**West, A.H. and Stock, A.M.** 2001. Histidine kinases and response regultor proteins in two-component signaling systems. Trends Biochem Sci. **26(6):** 369-376

**White, D.G., Goldman, J.D., Demple, B. and Levy, S.B.** 1997. Role of the *acrAB* locus in organic solvent tolerance mediated by expression of *marA, soxS,* or *robA* in *Escherichia coli.* J Bacteriol. **179(19):** 6122-6126

**Xiong, A., Gottman, A., Park, C., Baetens, M. Pandza, S. and Matin, A.** 2000. The EmrR protein represses the *Escherchia coli emrAB* multidrug resistance operon by directly binding to its promoter region. Antimicrob Agents Chemother. **44(10):** 2905-2907

**Xu, H., and Lee, S.F.** 2001 Activity of plant flavonoids against antibiotic resistant bacteria. Phytotherapy Research. **15:** 39-43

**Yamamoto, K., Hirao, K., Oshima, T. Aiba, H., Utsumi, R., and Ishihama, A.** 2005. Functional characterization *in vitro* of all two-component signal transduction systems in *Escherichia coli.* J Bio Chem. **280(2):** 1448-1456

**Yu, E.W., McDermot, G., Zgurskaya, H.I., Nikaido, H., and Koshland, D.E.** 2003. Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. Science. **300:**976-980

**Zheng, M., Dona, B., Schneider, T.D., and Storz, G.** 1999. OxyR and SoxRS regulation of *fur.* J Bacteriol. **181(15):** 4639-4643

**Zgurskaya, H.I. and Nikaido, H.** 1999. Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of Escherichia coli. Proc Natl Acad Sci. **96(13):** 7190-7195

**Zgurskaya, H.I. and Nikaido, H.** 2000. Multidrug resistance mechanisms: drug efflux across two membranes. Mol Microbiol. **37(2):** 219-225

**Zhou, H. Lowry, D.F., Swanson, R.V., Simon, M.I. and Dahlquist, F.W.** 1995. NMR studies of the phosphotransfer domain of the histidine kinase CheA from *Escherichia coli* : assignments, secondary structure, general fold, and backbone dynamics. Biochemisty. **34:** 13858-13870

**Zhou, L., Lei, X.H., Bochner, B.R. and Wanner, B.L.** 2003. Phenotype microarray analysis of *Escherichia coli* K-12 mutants with deletions of all two-component systems. J Bacteriol. **185(16):** 4956-4972